ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

PARTIALLY ORDERED ABELIAN GROUPS

BY A. H. CLIFFORD

(Received May 29, 1939)

The purpose of this paper is to point out a parallelism between certain aspects of the arithmetical theory of fields, as developed chiefly by Krull (1), and a new approach to the theory of function spaces recently developed by Kantorovitch (2). The unifying concept is that of a partially ordered abelian group. Theorems 1 and 2 below are concerned with the representation of such groups by means of vectors with components lying in linearly ordered abelian groups, while Theorem 3 deals with the completion of such groups. While writing this paper, I had the privilege of seeing the proofsheets of a beautiful paper by Lorenzen (3) on semigroups, which not only contains my results but goes considerably further. Proofs of the theorems will therefore be omitted, and I shall content myself with their interpretation as theorems on partially ordered groups in the additive notation. The procedure used in proving Theorem 3, for example, which is essentially the "v-Gruppensatz" of Artin, van der Waerden, and Krull (Krull, l.c., p. 120), is thus seen to be a direct generalization of the process of completing the additive group of rational numbers by means of Dedekind cuts.

In the classical theory of ideals in a finite algebraic field K, one succeeds in decomposing every principal ideal $(a) \neq (0)$ of K uniquely into a product of prime ideals:

$$(a) = \mathfrak{p}_1^{\alpha_1}\mathfrak{p}_2^{\alpha_2}\cdots\mathfrak{p}_n^{\alpha_n}.$$

Thus (a) can be represented by a vector $\{\alpha_1, \alpha_2, \cdots, \alpha_n, 0, 0, \cdots\}$ the components α_i of which are rational integers, all but a finite number of which are zero. The i^{th} component α_i of the vector is simply the power to which the i^{th} prime ideal $\mathfrak{p}_i$ occurs in the decomposition of (a). This vector representation has the following properties. *If*

$$(a) \rightarrow \{\alpha_1, \alpha_2, \cdots\}, \qquad (b) \rightarrow \{\beta_1, \beta_2, \cdots\},$$

$$(a + b) \rightarrow \{\gamma_1, \gamma_2, \cdots\},$$

then 1) $(ab) \rightarrow \{\alpha_1 + \beta_1, \alpha_2 + \beta_2, \cdots\}$,

2) *a divides b if and only if* $\alpha_i \leqq \beta_i$ $(i = 1, 2, \cdots)$,

and 3) $\gamma_i \geqq \min(\alpha_i, \beta_i)$.

In Krull's general (exponential) valuation theory, one simply removes the finiteness restriction and at the same time allows the exponents α_i to be arbitrary real numbers, or even to be elements of non-Archimedean linearly ordered

abelian groups (definition below). One then refers to the classical case as "finite and discrete." Since there is no reason to assume that the vector-components are denumerable or ordered in any way, Krull uses the notation $\{\cdots, \alpha_r, \cdots\}$.

Now the non-zero elements of K form a group G under multiplication, and a great many of the important concepts of multiplicative arithmetic can be formulated in G, the operation of addition not being used. On the one hand this leads to a new and more general theory—the arithmetic of abelian groups—and on the other hand (because of the simplification) it may lead back to new results in K. The latter assertion is borne out by Lorenzen's paper. One must of course assume the existence of a set S of "integers" in G such that 1) the product of two integers is an integer, 2) the identity element of G is an integer, and 3) every element of G is the quotient of two integers. We then define division in the usual way: a divides b if b/a is in S. The principal ideal (a) generated by a is defined to be the set of all elements of G divisible by a.

We may also begin with such a set S (called a *semigroup*), namely a set in which a multiplication ab is defined satisfying the associative and commutative laws, such that an identity element 1 exists in S, and such that the cancellation law (4) holds: $ac = bc$ implies $a = b$. S may then be extended to a group G in the usual way by means of formal quotients a/b. G is called the *quotient-group* of S.

G (or S) is said to be *reduced* if two elements which divide each other are necessarily equal. If G is not already reduced we can replace it by the group $\bar{G}$ of principal ideals in G, which has of course the same arithmetical structure as G itself. An element u of G is a *unit* if both u and u^{-1} are in S. Evidently $\bar{G}$ is isomorphic with the factor group G/U where U is the group of units of G.

We may now proceed to consider possible representations of the principal ideals of G by means of vectors, as described above for a field K. Condition 3) has now no meaning as it stands, since addition is not defined in G, so I shall simply discard it. (Actually Lorenzen retains 3) by referring it to a given system of ideals in G: if c is in the r-ideal $(a, b)_r$ then $\gamma_r \geqq \min (\alpha_r, \beta_r)$. For $r = s$ this is automatically satisfied. Theorem 2 below gives a representation satisfying this condition for $r = t$.)

A vector representation of G can plainly be regarded as an embedding of $\bar{G}$ in the group Γ of all vectors of the type considered, such that the division relation in $\bar{G}$ is identical with the component-wise order relation in Γ. Since Γ is an additive group, it seems natural to write $\bar{G}$ additively to begin with. Let therefore G be a reduced group with integral part S; write it additively, and write $a \leqq b$ if a divides b relative to S. One easily verifies the following, which I take as the defining axioms for a *partially ordered abelian group*:

I. *G is an abelian group under $+$.*

II. *G is a directed partially ordered set under $\leqq$, that is*:

1) $a \leqq a$,

2) $a \leqq b$, $b \leqq a$ *imply* $a = b$,

3) $a \leqq b, b \leqq c$ *imply* $a \leqq c$,

4) *given* a *and* b, *there exists* c *in* G *such that* $a \leqq c$, $b \leqq c$.

III. $a \leqq b$ *implies* $a + c \leqq b + c$ *for all* c *in* G.

Axioms II 1, 2, 3) are the usual axioms for a partially ordered set; the theory of such sets has become quite popular recently (see e.g. Mac Neille (5)). As usual, one writes $a \geqq b$ for $b \leqq a$, and $a < b$ or $b > a$ for $a \leqq b$, $a \neq b$. If A is a subset of a partially ordered set P, an element b of P is an *upper bound* of A if $a \leqq b$ for all a in A, and a *least upper bound* if $b \leqq c$ for every upper bound c of A. P is called a *directed* set if II 4) holds (6); this condition simply requires that every pair of elements of P (and hence every finite subset of P) be bounded from above. Lower bound and greatest lower bound are defined dually. If every pair of elements of P (and hence every finite subset of P) has a least upper and greatest lower bound, P is called by Birkhoff (7) a *lattice*. If, moreover, every subset of P bounded from above has a least upper bound, then P is a *complete* lattice; the dual property is then a consequence. P is *linearly* ordered if exactly one of the relations $a < b$, $a = b$, $a > b$ holds for each pair of elements a, b of P. Such a set is evidently a lattice, but need not be complete.

Let now G be a partially ordered abelian group. One easily verifies the usual rules for manipulating $+$ and $<$, e.g.

$$a_1 \leqq a_2,\, b_1 \leqq b_2 \quad \text{imply} \quad a_1 + b_1 \leqq a_2 + b_2,$$
$$a < b \quad \text{implies} \quad a + c < b + c \text{ (all } c).$$

An element $a \geqq 0$ is called *positive*; the set S of all positive elements of G is a semigroup. Evidently $a \leqq b$ if and only if $b - a$ is in S, and hence the given order relation $\leqq$ is the same as division relative to S. By II 4) every element a of G is the difference of two positive elements. For if c be chosen such that $a \leqq c$, $0 \leqq c$ then

$$a = c - (c - a), \qquad c \geqq 0, \qquad c - a \geqq 0.$$

Hence G is the difference-group of S. Finally, by II 2), G is reduced. We thus see that the notion of a partially ordered abelian group is identical with that of the quotient-group of a reduced semigroup.

We may now specialize the type of the order relation in G as indicated above for general partially ordered sets. Thus a *linearly ordered abelian group* is one in which the order relation is linear; a *lattice-ordered* (*complete ordered*) *group* is one which is a lattice (complete lattice) under the order relation. Before proceeding to a discussion of these types, and the consideration of more general types, a few preliminary remarks are necessary.

Observe that $a \leqq b$ if and only if $-a \geqq -b$. Hence the mapping $x \to -x$ is not only a group automorphism of G, but also a "dual automorphism" of G regarded as an ordered set. There is thus a "principle of duality" in G. For example, the dual of II 4) is automatically established. Likewise, if every

pair of elements a, b of G has a least upper bound, denoted by sup (a, b), then every pair a, b has also a greatest lower bound given by

$$\inf (a, b) = -\sup (-a, -b).$$

Two partially ordered groups G and G' will be called *order-isomorphic* if there is a group isomorphism $x \leftrightarrow x'$ between them which preserves order: if $a \leftrightarrow a'$, $b \leftrightarrow b'$, then $a \leqq b$ if and only if $a' \leqq b'$.

If G' is a subgroup of a partially ordered group G, then the order relation in G obviously orders G', and all the axioms are satisfied except possibly II 4). If II 4) holds for G', I shall call G' an *ordered subgroup* of G. This is equivalent to requiring that every element of G' shall be the difference of two positive elements of G'. On the other hand, a given partially ordered group H will be said to be *order-embedded* in G if it is order-isomorphic to a subgroup G' of G. G' is then naturally an ordered subgroup of G.

The prime example of a linearly ordered abelian group is the additive group P of real numbers, with $\leqq$ having its usual significance. P and all of its subgroups satisfy the *Archimedean axiom*:

If $\alpha > 0$, $\beta > 0$ *then there exists a natural number* n *such that* $n\alpha > \beta$.

Conversely, every linearly ordered abelian group satisfying this axiom is order-isomorphic with a subgroup of P. As an example of a non-Archimedean linearly ordered group, let Γ be the additive group of all real n-dimensional vectors $\{\alpha_1, \cdots, \alpha_n\}$. We order Γ lexicographically, i.e.

$$\{\alpha_1, \cdots, \alpha_n\} > \{\beta_1, \cdots, \beta_n\}$$

if the first non-vanishing difference $\alpha_1 - \beta_1$, $\alpha_2 - \beta_2$, $\cdots$ is positive. One easily verifies that Γ is a linearly ordered group under this definition. If $\epsilon_1, \cdots, \epsilon_n$ are the unit vectors and $i > j$, then $n\epsilon_i < \epsilon_j$ for all n. This construction can be extended to infinite vectors, and Hahn (8) has shown that every linearly ordered group is a subgroup of such a group. Since this more or less disposes of linearly ordered groups, we turn our attention to the construction of non-linear groups from linear ones.

The most obvious procedure is to form the (unrestricted) direct sum Γ of a set of linearly ordered groups Γ_τ. Γ may be thought of as the set of all possible "vectors" $\alpha = \{ \cdots, \alpha_\tau, \cdots \}$ with one component provided for each Γ_τ, and α_τ ranging independently over Γ_τ. Two vectors are added in the usual way:

$$\{ \cdots, \alpha_\tau, \cdots \} + \{ \cdots, \beta_\tau, \cdots \} = \{ \cdots, \alpha_\tau + \beta_\tau, \cdots \}.$$

We now order Γ component-wise:

$$\{ \cdots, \alpha_\tau, \cdots \} \leqq \{ \cdots, \beta_\tau, \cdots \} \quad \text{if } \alpha_\tau \leqq \beta_\tau \text{ for all } \tau.$$

One readily verifies that the axioms are satisfied. Moreover, Γ will be non-linear if there is more than one Γ_τ. We shall call Γ a *vector-group*. Greek letters will be used for vector-groups and their elements.

Actually Γ is lattice-ordered. If

$$\alpha = \{ \cdots, \alpha_\tau, \cdots \}, \qquad \beta = \{ \cdots, \beta_\tau, \cdots \}$$

then

$$\sup(\alpha, \beta) = \{ \cdots, \max(\alpha_\tau, \beta_\tau), \cdots \},$$

$$\inf(\alpha, \beta) = \{ \cdots, \min(\alpha_\tau, \beta_\tau), \cdots \}.$$

It is customary to write max and min instead of sup and inf in a linearly ordered group, since max (α, β) is simply α or β, whichever is larger.

If each Γ_τ is Archimedean, we call Γ an *Archimedean vector-group*. In this case the components α_τ of the vector α can be regarded as real numbers, and if we denote by T the index class over which the descriptive index τ ranges, then α may be regarded as a single-valued function on T to the real numbers. On the other hand we may start with any set T and define Γ to be the set of all real functions $\alpha = \alpha(\tau)$ on T. Addition and order in Γ may be defined in the obvious way:

$$\gamma = \alpha + \beta \textit{ is the function } \gamma(\tau) = \alpha(\tau) + \beta(\tau),$$

$$\alpha \leqq \beta \textit{ if } \alpha(\tau) \leqq \beta(\tau), \textit{ all } \tau \textit{ in } T.$$

Every Archimedean vector-group is an ordered subgroup of such a group Γ. In Hahn's construction mentioned above, one orders the set T in any way and forms the additive group of all functions α such that the set of all τ for which $\alpha(\tau) \neq 0$ is well-ordered under the ordering in T. Such functions can then be ordered lexicographically.

The problem with which we originally began can now be expressed precisely as the order-embedding of a given partially ordered group G in a vector-group Γ. Our first theorem characterizes those groups G for which such an embedding is possible.

Theorem 1. *A partially ordered abelian group G can be order-embedded in a vector-group if and only if it satisfies the condition: if some natural multiple na of an element a of G is positive, then a itself is positive.*

This is Lorenzen's Satz 14 for the special case $r = s$. The necessity of the condition is obvious. The sufficiency can also be proved, without the use of ideals, by paralleling Krull's proof (l.c. pp. 110–111) that a domain of integrity is a "principal order" if and only if it is integrally closed. I had known Theorem 1 only as an analog of Krull's theorem, whereas Lorenzen's beautiful result contains both as the special cases $r = s$ and $r = d$. However proved, it requires transfinite induction.

From the standpoint of a logical development of the theory and construction of partially ordered groups, Theorem 1 characterizes the immense class of all possible ordered subgroups of all possible vector-groups.

We pass now to lattice-ordered groups. Such a group, in the multiplicative notation, is simply one in which every pair of elements has a greatest common divisor. These were first studied in full generality by Dedekind (9), and some

years later by Levi (10). The fundamental properties of such groups have been derived anew in the additive notation by Kantorovitch (l.c. §1) and Freudenthal (11), apparently without realizing the parallel to elementary arithmetic. For example, their decomposition

$$a = a_+ - a_-, \qquad a_+ = \sup(0, a), \qquad a_- = \sup(0, -a),$$

is simply the expression of a fraction "in lowest terms." As a new contribution, they show that $|a|$ defined by

$$|a| = a_+ + a_-$$

has the usual properties of an absolute value. These writers are, however, primarily interested in the case when G is a real linear space, i.e. admits the real field as multiplier domain with the usual rules for multiplying vectors by scalars. This is always the case in function spaces of any interest.

From the point of view of general lattice theory, a lattice-ordered group is rather special, since it is distributive and self-dual. The former was known to Dedekind, but Freudenthal's proof (l.c., p. 642) is much simpler. The self-duality has been mentioned above.

One readily verifies that a lattice-ordered group G satisfies the condition of Theorem 1, and hence can be order-embedded in a vector-group Γ. Now Γ is a lattice and so is G, and the order relation in G is that of Γ. But it does not follow necessarily that G is a sublattice of Γ in the usual sense, i.e. if a and b are in G then sup (a, b) relative to Γ may not be the same as sup (a, b) relative to G. The next theorem takes care of this situation.

THEOREM 2. *Any lattice-ordered abelian group G can be order-embedded in a vector-group Γ such that G is a sublattice of Γ.*

This is Lorenzen's Satz 11 (along with Satz 4). In order to construct such a Γ one must introduce what Lorenzen calls *t-ideals* in G, this notion due originally to Arnold (12). These ideals can be defined in any G, but if G is lattice-ordered a t-ideal can be defined as a subset $\mathfrak{a}$ of G bounded from below and such that

1) *if a is in $\mathfrak{a}$ and $a \leqq b$, then b is in $\mathfrak{a}$;*

2) *if a and b are in $\mathfrak{a}$, then* inf (a, b) *is in* $\mathfrak{a}$. Except for the boundedness condition, this (or its dual) is Stone's definition of an ideal in any distributive lattice (13).

(As a rather interesting minor theorem not mentioned by Lorenzen, it is possible to set up a one-to-one correspondence between the prime subideals of a given prime $\mathfrak{p}$ and the prime ideals of the corresponding valuation semigroup $S_\mathfrak{p}$ under which the inclusion relation is invariant. As a consequence, which can also be proved directly, if $\mathfrak{p}$ and $\mathfrak{q}$ are any two prime ideals then either they are relatively prime or one contains the other. Cf. Krull, l.c., p. 112.)

We come now to the consideration of complete ordered groups. (In Lorenzen's terminology, a lattice-ordered group is called "v-complete," or simply "complete," while a complete ordered group is called "totally complete." His

terminology is flawless, but I prefer to reserve the term "complete" for its well-established topological meaning. Kantorovitch calls it a "partially ordered topological group.") For this theory I can only refer to §3 of Kantorovitch's paper. Here a notion of convergence is introduced; the functions $x + y$, sup (x, y), inf (x, y), and $| x |$ are shown to be continuous, and the Cauchy convergence criterion is proved. I do not believe that the resulting topology has ever been used in arithmetic.

(I might remark here that the results in Kantorovitch §2 hold without the assumption that G admits the field of real numbers, and that Satz 21f in §4 holds in any lattice-ordered group. One simply uses standard arithmetical arguments, e.g. that if inf $(a, b) = 0$ then inf $(\lambda a, \mu b) = 0$ for arbitrary positive integers λ, μ.)

Our final theorem deals with the question of when a given partially ordered group can be completed.

THEOREM 3. *A partially ordered abelian group G can be order-embedded in a complete ordered group H if and only if G satisfies the condition: if the set of all natural multiples na of an element a of G is bounded from below, then a is positive.*

Moreover, H is essentially unique if we require that every element of H be the greatest lower bound of a set of elements of G.

The condition stated is what Krull calls "complete integral closure," and Lorenzen "total closure," and the sufficiency is an immediate consequence of the Artin-van der Waerden-Krull "v-Gruppensatz" (Krull, l.c., p. 120). The notion of a v-ideal can be introduced in any partially ordered set P: a subset $\mathfrak{a}$ of P bounded from below is a *v-ideal* if it contains every upper bound of the set of all lower bounds of $\mathfrak{a}$. The set of all v-ideals in P constitutes a complete lattice L under set-theoretical inclusion. P is order-embedded in L by identifying the element a of P with the principal ideal (a) it generates. Moreover, P is *normally* embedded in L, by which I mean that every element of L is the greatest lower bound of a subset of P. This means of completing a partially ordered set was used by Mac Neille (l.c., p. 443); indeed, a v-ideal is the same thing as the upper class of a cut. Mac Neille's definition of a cut is a direct generalization of that of a Dedekind cut in the rational numbers.

If G is a partially ordered *group*, then we can add two cuts or v-ideals $\mathfrak{a}$, $\mathfrak{b}$ in exactly the same way that we add two Dedekind cuts: $\mathfrak{a} + \mathfrak{b}$ is defined to be the smallest v-ideal containing all the sums $a + b$ with a in $\mathfrak{a}$, b in $\mathfrak{b}$. Addition so defined coincides with that in G, since for principal ideals we have $(a) + (b) = (a + b)$, and satisfies all the postulates for an abelian group *except the existence of negatives*. In fact, the set $\mathfrak{G}$ of all v-ideals in G satisfies all our postulates for a complete ordered group except for this one lack. But the v-Gruppensatz states that $\mathfrak{G}$ *will* be a group if and only if G is completely integrally closed.

The necessity of this condition is almost immediate. It holds for any complete ordered group (see e.g. Kantorovitch, Satz 15f), and hence for any ordered subgroup thereof.

The second part of the theorem is an obvious extension of the unicity theorem

of Arnold (l.c.) in the finite discrete case. Because of the normality condition one is able to set up a one-to-one correspondence $\alpha \leftrightarrow \mathfrak{a}$ between H and the group $\mathfrak{G}$ of v-ideals in G such that 1) $\mathfrak{a}$ is the set of all upper bounds of α in G, and 2) α is the greatest lower bound of $\mathfrak{a}$ in H. This correspondence is an order-isomorphism between H and $\mathfrak{G}$ leaving the elements of G invariant. This unicity theorem does not hold for arbitrary partially ordered sets.

The v-system plays the decisive rôle in the completion process because *it is the only system of ideals which can possibly form a group.* It is the only system which can do the job, does it well if it is possible at all, and then expires nobly—for in a complete group every v-ideal is principal and hence of no further use. For further study one must use a weaker system, preferably the t-system, or avail oneself of the Kantorovitch topology.

Regarding the condition of "complete integral closure" or "total closure" occurring in Theorem 3, it is easily seen that it is equivalent to the Archimedean axiom in the case of a linearly ordered group. (This shows incidentally that a non-Archimedean linear group cannot be completed, at least in the sense here used.) We might therefore adopt it as the defining property of an *Archimedean partially ordered group.* In fact, Kantorovitch refers to it as the "Archimedean axiom" in an earlier paper.

One easily sees that an Archimedean vector-group, as previously defined, and hence any ordered subgroup thereof, is Archimedean in this sense. Whether or not the converse is true—that every Archimedean group can be order-embedded in an Archimedean vector-group—is still an open question. This is a conjecture of Krull; Lorenzen proves a special case (§5, l.c.). I have been working on this for the past year, for the most part using t-ideals, but with no success whatever. Perhaps the Kantorovitch topology may open up a clue to the mystery.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
CAMBRIDGE, MASS.

REFERENCES

(1) W. Krull, *Idealtheorie* (Ergebnisse der Math. vol. 4, 1935, no. 3), §§5, 6.

(2) L. V. Kantorovitch, *Lineare halbgeordnete Räume* (Mat. Sbornik 2, 1937, 121–165).

(3) P. Lorenzen, *Abstrakte Begründung der multiplikativen Idealtheorie* (Math. Zeit. 45, 1939, 533–553.)

(4) In a previous paper (Annals of Math. 39, 1938, 594–610) I discussed the arithmetical theory of semigroups for which this condition was not postulated; those for which it holds were called "regular."

(5) H. M. Mac Neille, *Partially ordered sets* (Trans. A. M. S. 42, 1937, 416–460).

(6) This term used by Garrett Birkhoff (Annals of Math. 38, 1937, p. 40). E. H. Moore and H. L. Smith, to whom the notion is due, call it the "composition property" (Amer. Jour. 44, 1922, p. 103).

(7) Garrett Birkhoff, *On the combination of subalgebras* (Proc. Cambr. Phil. Soc. 29, 1933, 441–464); his new book on lattices should appear shortly.

(8) H. Hahn, *Über die nichtarchimedischen Grössensysteme* (S. B. d. Math.-Nat. Klasse d. Wiener Akad. 116IIa, 1907, 601–655).

(9) R. Dedekind, *Über Zerlegungen von Zahlen durch ihre grössten gemeinschaftlichen Teiler* (Ges. Werke, Bd. 2, XXVIII, §6).

(10) F. Levi, *Arithmetische Gesetze in Gebiete diskreter Gruppen* (Rend. Palermo 35, 1913, 225-236).

(11) H. Freudenthal, *Teilweise geordnete Moduln* (Proc. Akad. Wet. Amsterdam 39, 1936, 641-651).

(12) J. Arnold, *Ideale in kommutativen Halbgruppen* (Rec. Math. Moscou 36, 1929, 401-407).

(13) M. H. Stone, *Topological representations of distributive lattices and Brouwerian logics* (Čas. Mat. a Fys. 67, 1937, 1-25).

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

QUASI-GROUPS[1]

BY G. N. GARRISON

(Received June 3, 1939)

Introduction

Briefly stated, a *quasi-group* is a multiplicative system in which, given any two of the three elements a, b, c, the third is uniquely determined by the relation $ab = c$. Obviously, multiplication is not necessarily associative. For this reason the quasi-group may be used to study generalizations of the associative law. Such studies have been made by A. Suschkewitsch,[1a] B. A. Hausmann and O. Ore,[2] and, more recently, by D. C. Murdoch.[3]

A different point of view is taken in the present paper. No form of associative law is required for the entire quasi-group. Instead, we have endeavored to discover properties of subsets of its elements.

Section 1 is devoted to fundamental definitions. It also contains all the formal properties of complexes which we shall require. Section 2 deals with subsets of elements which have special associativity properties. Some of its results have been stated previously without proof by E. Schönhardt[4] for quasi-groups which contain a two-sided identity element. In Section 3 we correlate the results of the previous section for the case of commutative multiplication.

Finally, in Section 4 we discuss quotient quasi-groups and homomorphisms of a quasi-group. This theory depends upon our notion of an invariant complex H (which is not necessarily a sub-quasi-group) defined by relations of the form

$$(Ha)(Hb) = Hc.$$

We are able to show that every complex with this property determines a homomorphism and conversely, that every homomorphism of a quasi-group is isomorphic with the quotient quasi-group determined by a suitably chosen invariant complex.

I would like to acknowledge my appreciation of the privilege I have had of discussing the details of this paper with Professor J. H. M. Wedderburn.

---

[1] Presented to the American Mathematical Society, Dec. 26, 1939.

[1a] A. Suschkewitsch, *On a generalization of the associative law*, Trans. of the Amer. Math. Soc., vol. 31 (1929), pp. 204–214.

[2] B. A. Hausmann and O. Ore, *Theory of Quasi-groups*, Amer. Jour. of Math., vol. 59 (1937), pp. 983–1004.

[3] D. C. Murdoch, *Quasi-groups which satisfy certain generalized associative laws*, Amer. Jour. of Math., vol. 61 (1939), pp. 509–522.

[4] E. Schönhardt, *Über lateinische Quadrate und Unionen*, Jour. für Math., vol. 163 (1930), p. 203.

1. The quasi-group

1.1. A finite set S of n elements $a_1, a_2, \cdots, a_n$ is called a *quasi-group* of order n when:

i) S possesses the group property, that is, when each of the n^2 products $a_i a_j$ designates one and only one element of S; and

ii) $a_k a_i = a_k a_j$ and $a_i a_k = a_j a_k$ only if $i = j$ for all k from $1, 2, \cdots, n$.

We have at once the following two theorems:

THEOREM 1.1. *Condition* (ii) *implies that there exists in S a unique x satisfying $ax = b$ and a unique y satisfying $ya = b$ for each pair of values of a, b in S.*

For by (ii) the n elements aa_j $(j = 1, 2, \cdots, n)$ are all distinct. Thus one of them is the element b. Let aa_p denote it. Then $x = a_p$ is the required solution. The other part of the theorem follows similarly.

THEOREM 1.2. *A sufficient condition that a quasi-group be a group is that $a_i(a_j a_k) = (a_i a_j)a_k$ $(i, j, k = 1, 2, \cdots, n)$.*

For, condition (i), the result of Theorem 1.1, and the hypothesis of the present theorem, form the definition of a group used by some writers. It has been shown to be equivalent to the more usual one in terms of an identity element and inverses.[5]

1.2. Any sub-set H of the elements of a quasi-group S will be called a *complex* H in S; a complex that does not contain all the elements of S is called a *proper complex.*

The complex composed of those elements common to two or more complexes $H, K, M, \cdots$ will be denoted by $H \wedge K \wedge M \wedge \cdots$ and will be called their *intersection.* If $H \wedge K$ be a null-set we write $H \wedge K = 0$. To denote the complex consisting of those elements which are contained in at least one of the complexes we shall write $H + K + M + \cdots$ and shall call it the *sum* of $H, K, M, \cdots$.

By the *product* HK of two complexes H and K we shall mean that complex which contains all possible distinct products of the form hk where h and k denote elements of H and K respectively. And by $H = K$ we mean that H and K consist of exactly the same set of elements of S.

1.3. The number of distinct elements contained in a given complex is its *order.* We now list a number of simple theorems to which we shall have frequent occasion to refer.

Let H, K, M, and T denote complexes in a quasi-group S. Then we have:

i) *If H be of order h, so are also Hs and sH for any element s in S. If also K be of order k, then HK (or KH) is of order at least as great as the larger of h and k.*

ii) $s(H \wedge K) = sH \wedge sK$ *and* $(H \wedge K)s = Hs \wedge Ks$.

For if a be any element of both H and K, then sa is an element of both sH

[5] Hasse, H., *Hoehere Algebra*, p. 52.

and sK. Hence

$$s(H \wedge K) \leqq sH \wedge sK. \tag{1}$$

Also if b be any element of both sH and sK then $b = sh = sk$ for some pair $h < H$, $k < K$. By §1.1 (ii) $h = k$ and $b < s(H \wedge K)$. Thus

$$sH \wedge sK \leqq s(H \wedge K). \tag{2}$$

We compare (1) with (2) and have the theorem.

iii) *If $Hs = Ks$ or $sH = sK$ for some s, then $H = K$.*

For if h be an element of H, then for some $k < K$ we have $hs = ks$. Thus $h = k$ (§1.1 ii) and $H \leqq K$. Similarly $K \leqq H$. Hence the theorem.

iv) *If H, K, and M all three have the same order, and if $HK = M$, then also $Hk = M$ and $hK = M$ for all k in K and h in H.*

For we have Hk the same order as H (cf. (i) above) and thus the same order as M. But all elements of Hk lie in HK and thus in M, so that $Hk = M$.

v) *If H, K, M, and T all have the same order, and if $HT = M = KT$ or $TH = M = TK$, then $H = K$.*

For if $t < T$, $Ht = M = Kt$ by (iv) and $H = K$ by (iii).

1.4. By a *sub-quasi-group* H we mean a complex H in S whose elements form according to §1.1 a quasi-group. We make three comments which will be useful later.

i) *A complex H is a sub-quasi-group of S if and only if $HH = H$.* The condition is necessary to satisfy §1.1 (i); and it is sufficient since it satisfies §1.1 (i) and since $h_k h_i = h_k h_j$ for $i \neq j$, being impossible in S, is also in H, so that §1.1 (ii) is satisfied.

ii) *The order of H need not divide the order of S.*

This is seen from the example below where $H = (a, b)$ is a sub-quasi-group of order 2.

$$\begin{array}{c|ccccc} & a & b & c & d & e \\ \hline a & a & b & d & e & c \\ b & b & a & e & c & d \\ c & e & d & c & b & a \\ d & c & e & a & d & b \\ e & d & c & b & a & e \end{array}$$

iii) *The intersection of two sub-quasi-groups is either a sub-quasi-group or is zero.*

For the example in (ii) shows that the intersection may be zero, for there $(a, b) \wedge (e) = 0$. If now $H \wedge K = T \neq 0$, then tt' lies in both K and H, and thus in their intersection T, so that $TT = T$.

2. Center, left, and right associative elements

2.1. We shall call an element c of S *center associative* in S if $x(cy) = (xc)y$ for all x, y. We then have the following

THEOREM 2.1. *The complex C consisting of the totality of center associative elements of a given quasi-group S is a group. Its identity element e is both a left and a right identity element for S itself. Furthermore, if for c in C, s in S, we have $cy = s$, and $wc = s$, then $y = c^{-1}s$, and $w = sc^{-1}$. Finally, the order of C divides the order of S.*

The proof will be made by establishing five lemmas.

LEMMA 1. *The complex C is a sub-quasi-group, that is, $CC = C$.*

To show this, let c, d be a pair of elements in C. We shall prove that cd is a center associative element which shows that it lies in C since C is the totality of such elements. We have

$$(3) \qquad (xc)d = x(cd), \qquad c(dy) = (cd)y$$

for all x, y since c and d are center associative. Since d is center associative we have from (3)

$$(4) \qquad (xc)(dy) = ((xc)d)y = (x(cd))y.$$

Also since c is center associative we see that

$$(5) \qquad (xc)(dy) = x(c(dy)) = x((cd)y).$$

We now compare (4) with (5) and find that cd is a center associative element. This proves the lemma.

LEMMA 2. *The sub-quasi-group C is associative, and is thus by Theorem 1.2 a group.*

For, since $x(cy) = (xc)y$ for all x, y in S, we have for all c, d, f in C the relation: $d(cf) = (dc)f$. Hence Theorem 1.2 applies.

LEMMA 3. *The identity element e of C is both a left and a right identity element for S.*

PROOF. Let c be an element of C, and let x vary over S. Then cx varies over S by §1.1 (ii). Since $e(cx) = (ec)x = cx$ for all x we see that e is a left identity in S. Similarly, e is a right identity.

LEMMA 4. *The equations $cy = s$, $wc = s$ for c in C and s in S have the unique solutions $y = c^{-1}s$, $w = sc^{-1}$ where c^{-1} denotes the inverse of c in C.*

For unique solutions exist by Theorem 1.1. We verify the above expressions for them by substitution; thus

$$cy = c(c^{-1}s) = (cc^{-1})s = es = s$$

since c^{-1}, being in C, is center associative and since Lemma 3 applies here. The other solution is checked similarly.

LEMMA 5. *The order k of C divides the order n of S.*

PROOF. We shall show that S may be represented as a sum of non-intersecting complexes Cs_i, each of order k, in such a way that each element of S lies in one, and only one, such complex.

Let s_1 be any element of S not contained in C. Then Cs_1 contains k distinct elements of S. Also $Cs_1 \wedge C = 0$, for if not we would have $cs_1 = c'$ for c, c'

in C, from which follows $s_1 = c^{-1}c'$ by Lemma 4; but this contradicts the hypothesis that s_1 be not in C.

Next let s_2 be an element of S not contained in either C or Cs_1. Then Cs_2 is of order k with $Cs_2 \wedge C = 0$. We have also $Cs_2 \wedge Cs_1 = 0$ since $cs_2 = c's_1$ implies $s_2 = (c^{-1}c')s_1$ in view of the center associativity of c'. Thus we reach another contradiction, namely, s_2 an element of Cs_1.

Proceeding in this way we show that it is possible to select elements s_1, $s_2, \cdots, s_t$ from S such that

$$S = Cs_1 + Cs_2 + \cdots + Cs_t$$

with $Cs_i \wedge Cs_j = 0$ for $i \neq j$. This proves the lemma and completes the proof of Theorem 2.1.

COROLLARY. *We may also represent S as a sum of complexes of the form s_iC.*

2.2. *Left and right associative elements.* By a *left associative* element m we mean one which satisfies the relation $m(xy) = (mx)y$ for all x, y; and by a *right associative* element r one which satisfies $(xy)r = x(yr)$ for all x, y.

We now state a theorem similar to Theorem 2.1 which, though similar, still differs in enough ways to make a precise statement desirable.

THEOREM 2.2. *The complex L (or R) consisting of the totality of left (right) associative elements of a given quasi-group is a group. The identity element e of $L(R)$ is a left (right) identity of S itself. The equation $mx = s$, m in L, s in S, has the unique solution $x = m^{-1}s$, while $yr = s$, r in R, has the unique solution $y = sr^{-1}$. Also, the order of L (or R) divides the order of S.*

PROOF. Much of this proof need not be written out in detail since obvious modifications of Lemmas 1–5 will follow here. Certainly nothing need be carried through for both L and R. We shall however give the detailed proof that $LL = L$. Thus, let m, n be a pair of elements in L, and x, y any pair in S. Then

$$m((nx)y) = m(n(xy)) = (mn)(xy)$$
$$= (m(nx))y = ((mn)x)y$$

from which we see that mn is left associative and is thus contained in L. Finally, in order to apply the methods of Lemma 5 to prove the last part of the theorem, we must be careful to form complexes $Ls_i(s_jR)$ so that we have:

$$S = Ls_1 + Ls_2 + \cdots + Ls_i$$

or

$$S = s_1'R + s_2'R + \cdots + s_j'R.$$

We turn now to another interesting complex which is a group.

THEOREM 2.3. *Let M be a sub-quasi-group of S, such that any pair of elements m, m' in M satisfy the relation*

$$m(m'x) = (mm')x$$

for all x. Then M is a group. Its identity e is a left identity of S. The equation $mx = s$ $(m < M, s < S)$ is solved by $x = m^{-1}s$. Finally, its order divides the order of S.

PROOF. Since $(mm')m_1 = m(m'm_1)$ we have by Theorem 1.2 that M is a group. Let e denote its identity element. Then $m(ey) = (me)y = my$ so that $ey = y$ for all y. Thus e is a left identity in S. The relation $m(m^{-1}s) = (mm^{-1})s = es = s$ verifies $m^{-1}s$ as the solution x of the equation $mx = s$. The proof that the order of M divides the order of S is similar to the proof of Lemma 5.

As examples of groups satisfying the conditions for M we mention the groups L and C already discussed.

THEOREM 2.4. *When any two of the four groups C, L, R, M exist in S their identity elements correspond.*

For if there were two left identities e and e' we would have $ex = x = e'x$ which is impossible by §1.1 (ii). Similarly there is only one right identity f. But when both of these exist they are the same since $e = ef = f$.

2.3. The following theorems are of some interest in connection with the classification of quasi-groups.

THEOREM 2.5. *A necessary condition for the existence, in a quasi-group S, of L, (or M), R, or C is the existence of a left identity, a right identity, or both, respectively.*

THEOREM 2.6. *A left identity e is left associative, a right identity is right associative, and an element which is both is left, right, and center associative.*

For if $ex = x$ for all x, then $(ex)y = xy = e(xy)$, so that e is left associative. The rest of the proof is similar.

As a final remark, we observe that the intersections of any two or more of C, L, R, M are groups. For their intersections are never vacuous (Theorem 2.4). Hence by §1.4 (iii) they are sub-quasi-groups. And finally, by argument similar to the proof of Lemma 2, they are groups.

3. Commutativity and normal complexes

3.1. An element c will be called *commutative* in S when $cx = xc$ for all x in S. If every element of S is commutative in S, then S will be called commutative.

The results of Section 2, for S commutative, are related by the following

THEOREM 3.1. *In a commutative quasi-group a left associative element is right associative and conversely; furthermore, such elements are also center associative;*[6] *that is, $L = R \leqq C$.*

PROOF. Let b be left associative. Then $b(xy) = (bx)y$, hence $(xy)b = y(bx)$ and thus $(yx)b = y(xb)$ which proves the first part of the theorem if the steps be read both as written and backwards. To prove b is also center associative we write $(bx)y = (xb)y$ and compare with $(bx)y = b(xy) = b(yx) = (by)x = x(by)$.

[6] I have not been able to prove or to disprove the converse.

3.2. A complex H contained in S will be called *normal* in S if for all x we have $Hx = xH$. It is clear that the following complexes are normal in S:

i) every complex, if S be commutative;
ii) any complex whose elements are commutative in S;
iii) the complex S itself;
iv) the intersection of two normal complexes if it is not a null-set.

3.3. We now prove a theorem which demonstrates the existence of a type of complex which we shall study closely later.

THEOREM 3.2. *Let H be a sub-group of any two of the groups C, L, or R (§§2.1–2.3) and let H be normal in S. Then for any pair of elements x, y in S we have $(Hx)(Hy) = H(xy)$; and also, H being normal, $(xH)(yH) = (xy)H$.*

PROOF. We give the detailed proof for the case $H \leqq (C \wedge L)$. The others are all made similarly. Let small letters h_i denote elements of H. Then

$$\begin{aligned}(h_1x)(h_2y) &= h_1(x(h_2y)) \text{ since } h_1 \text{ is left associative;}\\ &= h_1((xh_2)y) \text{ since } h_2 \text{ is center associative;}\\ &= h_1((h_3x)y) \text{ since } H \text{ is normal;}\\ &= h_1(h_3(xy)) \text{ since } h_3 \text{ is left associative;}\\ &= (h_1h_3)(xy) \text{ since } h_1 \text{ is left associative.}\end{aligned}$$

This means that the product $(Hx)(Hy) \leqq H(xy)$ but by §1.3 (i) the inclusion sign is not possible. Hence the theorem follows.

4. Invariant complexes and quotient quasi-groups

4.1. A complex H is called *invariant* in S if, for all a, b in S there exists a corresponding element c such that $(Ha)(Hb) = Hc$; and (as we shall prove in Theorem 4.9 and below) an element d such that $(aH)(bH) = dH$.

We observe at once that such complexes exist in view of Theorem 3.2 in which we showed the existence of a complex H with properties more special than our definition requires, for there $Hc = H(ab)$ and $a < Ha$ for all a were also true. We shall consider these cases later.

Until after the proof of Theorem 4.9 we shall distinguish between the two cases of our definition by using the terms left and right invariant. We turn now to a study of the properties of a left invariant complex in an arbitrary quasi-group.

THEOREM 4.1. *If H be left invariant and if $(Ha)(Hb) = (Ha)(Hc)$, then $Hb = Hc$; and if $(Hb)(Ha) = (Hc)(Ha)$, then also $Hb = Hc$.*

For since Ha, Hb, Hc are all of the same order by §1.3 (i), the theorem follows by §1.3 (v).

THEOREM 4.2. *Each element b of S lies in some complex Hx.*

PROOF. Form Hx_i for all x_i in S. We get thereby $HS = S$ by §1.3 (i), so that b cannot be omitted from every complex.

THEOREM 4.3. *For any p in S, $(Ha)p$ is some complex Hb.*

For we saw in Theorem 4.2 that p lies in some complex Hm. Then by §1.3 (iv), $(Ha)p = (Ha)(Hm) = Hb$.

THEOREM 4.4. *Let a, b denote any pair of elements of S. Then either $Ha \wedge Hb = 0$ or $Ha = Hb$.*

PROOF. Assume $Ha \neq Hb$ so that two cases are possible: (i) $Ha \wedge Hb = 0$, and the theorem is proved; (ii) $Ha \wedge Hb = T \neq 0$, T proper in Ha and Hb. In this case we have $(Ha)T \leqq (Ha)(Ha)$ (i.e., $\leqq Hc$ for some c). Now by §1.3 (i) the order of $(Ha)T$ is not less than the order of Ha but from the previous step the order of $(Ha)T$ is not greater than the order of Hc. Since Ha and Hc are of the same order we have $(Ha)T = (Ha)(Ha)$; similarly $(Ha)T = (Ha)(Hb)$, and thus by Theorem 4.1, $Ha = Hb$.

THEOREM 4.5. *The elements of H itself all belong to some single complex Ha.*

PROOF. If the theorem were not true, then for some h and h' in H we would have $h < Ha$, $h' < Hb$, $Ha \wedge Hb = 0$ for some a, b. Let now k be any element of S. Then $hk < (Ha)k$, $h'k < (Hb)k$ with $(Ha)k \wedge (Hb)k = 0$ by §1.3 (ii). Now by Theorem 4.3, $(Ha)k$ and $(Hb)k$ are complexes of the form Hp, Hq. We have then shown that $hk < Hp$ and $h'k < Hq$ where $Hp \wedge Hq = 0$. When we consider that both hk and $h'k$ lie in Hk we reach a contradiction of Theorem 4.4.

In regard to the relation existing between the set of complexes Hb and the elements b which determine them we have

THEOREM 4.6. *If $Ha = Hb$, then a and b both lie in the same complex Hc, and conversely.*

PROOF. Assume $a < Hc$ and $b < Hd$. Then $Ha = H(Hc)$ and $Hb = H(Hd)$ (Theorem 4.5); hence $H(Hc) = H(Hd)$ so that $Hc = Hd$ by Theorem 4.1. Conversely, if $a < Hc$ and $b < Hc$, then $Ha = H(Hc) = Hb$.

It is now clear, in view of Theorems 4.4–4.6, that elements $a_1, a_2, \cdots, a_r$ exist in S such that

$$S = Ha_1 + Ha_2 + \cdots + Ha_r \qquad (Ha_i \wedge Ha_j = 0,\ i \neq j), \tag{6}$$

in which we may without loss of generality take $Ha_1 = H$. Also observe that the decomposition (6) is not altered if any a_i be replaced by any other element of S which lies in the same complex Hx with it (Theorem 4.6).

THEOREM 4.7. *The set of complexes Ha_i of (6) forms a quasi-group with product of complexes as the "product" of §1.1 (i).*

For the product of two complexes is a complex since H is left invariant; thus §1.1 (i) is satisfied. The result of Theorem 4.1 satisfies §1.1 (ii).

DEFINITION. The group whose existence is proved in Theorem 4.7 is called the quotient quasi-group of S by H and is denoted by S/H (tentatively by $H^{-1}S$ until after Theorem 4.9).

THEOREM 4.8. *Let H be an invariant complex. Then for any a in S, Ha is also an invariant complex.*

PROOF. Consider the product $((Ha)b)((Ha)c)$ which by Theorem 4.3 takes the form $(Hp)(Hq) = Hr$. Now in $H^{-1}S$ we solve the equation (cf. Theorem 1.1) $(Ha)X = Hr$ for the complex X. Let d be an element of X. Then $Hr = (Ha)d$. We then see from the discussion above that $((Ha)b)((Ha)c) = (Ha)d$ so that Ha is invariant.

We are now in a position to show that a left invariant complex is also right invariant and thus to justify the parenthetical remark in our definition of an invariant complex.

THEOREM 4.9. *If b is any element of S, the complex bH is also of the form Hp for some p in S, and conversely. Thus a left invariant complex is right invariant and conversely.*

For by Theorem 4.2, b lies in some complex Hx; then by §1.3 (iv), $bH = (Hx)H$, so that by Theorems 4.5 and 4.7, $bH = Hp$ where $Hp = (Hx)H$ in $H^{-1}S$.

Conversely, Hp is of the form bH since in $H^{-1}S$ we can solve the equation $Hp = XH$ for the complex X by Theorem 1.1. Let b be any element in X; then $Hp = bH$ by §1.3 (iv).

It is now clear that $(bH)(cH) = (Hp)(Hq) = Hr = dH$ which proves the theorem.

Obvious analogues of Theorems 4.1–4.6 show that a decomposition of S of the form

$$S = b_1H + b_2H + \cdots + b_rH \tag{7}$$

analogous to (6) is possible. Clearly, the b_iH of (7) are the same as the Ha_i of (6) in some order in view of Theorems 4.9 and 4.4.

We conclude this section with three theorems on quotient quasi-groups.

THEOREM 4.10. *Let K be a sub-quasi-group of a quasi-group S, and let H be an invariant complex contained in K. Then H is invariant in K, and K/H is a sub-quasi-group of S/H.*

PROOF. Since $H < K$, the complexes Hk_1, Hk_2 consist of elements of K; thus their product $(Hk_1)(Hk_2)$ consists of elements of K. Since now H is invariant in S, there exists an element c in S such that $(Hk_1)(Hk_2) = Hc$. We shall prove that c lies in K. Suppose this were not so. Then $K + c$ is a complex of greater order than K. Then $H(K + c)$ is of greater order than K by §1.3 (i). This contradicts the fact that $H(K + c) \leqq HK + Hc \leqq K$ since $HK \leqq K$ and $Hc \leqq K$ from the preceding discussion. Thus H is invariant in K, K/H exists and forms a sub-quasi-group of S/H by §1.4 (i).

THEOREM 4.11. *Let Q be a sub-quasi-group of S/H. Then the set Q' of elements of S belonging to the complexes which lie in Q forms a sub-quasi-group of S.*

PROOF. Let Q consist of complexes $Hq_1, Hq_2, \cdots, Hq_t$ where $q_1, q_2, \cdots, q_t$ are a subset of the a_i in (6). Then Q' is the complex $Hq_1 + \cdots + Hq_t$ in S. Let a', b' be elements of Q'. Then for some i, j we have $a' < Hq_i$, $b' < Hq_j$ and $a'b' < (Hq_i)(Hq_j) = Hq_k$ since $QQ = Q$. Thus $a'b' < Q'$, and Q' is a sub-quasi-group by §1.4 (i).

THEOREM 4.12. *Let Q be an invariant complex in S/H. Then the set Q' of elements of S belonging to the complexes which are contained in Q forms an invariant complex in S.*

PROOF. If Q be invariant in S/H, then a complex Ha_k exists in S/H such that $(Q(Ha_i))(Q(Ha_j)) = Q(Ha_k)$. If in S, $a' < Ha_i$, then

$$\begin{aligned} Q'a' &= (Hq_1 + \cdots + Hq_t)a' \\ &= (Hq_1)a' + \cdots + (Hq_t)a' \\ &= (Hq_1)(Ha_i) + \cdots + (Hq_t)(Ha_i) \text{ by §1.3 (iv)} \\ &= Q(Ha_i). \end{aligned}$$

If also $b' < Ha_j$ and $c' < Ha_k$, then by repeating the argument just made we have $(Q'a')(Q'b') = Q'c'$. Thus Q' is invariant in S.

4.2. *Homomorphism of a quasi-group.* The quotient quasi-group defined in §4.1 is readily seen to be an example of a system $\bar{S}$ satisfying the following

DEFINITION. Let $\bar{S}$ denote a set of elements $\bar{a}_1, \bar{a}_2, \cdots$ satisfying the closure property (§1.1 i) of a quasi-group with respect to some operation. Let S be any quasi-group such that there exists a correspondence $a \to \bar{a}$ from S to $\bar{S}$ with the two properties:

i) to each a in S there corresponds one and only one $\bar{a}$ in $\bar{S}$;

ii) no $\bar{a}$ in $\bar{S}$ is without a correspondent in S.

We call the set $\bar{S}$ a *homomorph* of S when the correspondence satisfies

iii) if $a \to \bar{a}$, $b \to \bar{b}$, then $ab \to \bar{a}\bar{b}$.

If the correspondence is one-to-one $\bar{S}$ is called an *isomorph* of S and S and $\bar{S}$ are said to be *isomorphic.*

In the discussion of homomorphism of a quasi-group we shall need the following

LEMMA 6. *Let $H_1, \cdots, H_t$ be a set of complexes in S such that*

$$H_i \wedge H_j = 0 \qquad (i \neq j) \tag{8}$$

and such that for all i, j we have

$$H_iH_j \leqq H_k \tag{9}$$

for some k. Then the complexes H_i are all of the same order, and $H_iH_j = H_k$.

PROOF. We assume the lemma false. Then without loss of generality we may take $H_1, \cdots, H_r$ to be of maximal order d, and $H_{r+1}, \cdots, H_t$ to be of order less than d. Consider now the set of complexes H_1H_i, $i = 1, \cdots, r$. By §1.3 (i) their orders are not less than d, and thus by (9), since d is maximal, their orders are d. They are distinct since $H_1H_i = H_1H_j$ implies $H_i = H_j$ by §1.3 (v). Thus they are $H_1, \cdots, H_r$ in some order. Consider now H_1H_{r+1}, which by §1.3 (i) is also of order d. This means, in view of (9), that

$$H_1H_{r+1} = H_1H_j \tag{10}$$

for some j from $1, \cdots, r$. Now let h be an element of H_1. Then by §1.3 (iv) we have $hH_j = H_1H_j$ if $1 \leqq j \leqq r$. Thus from (10) $hH_{r+1} \leqq hH_j$ which implies that $hH_{r+1} \wedge hH_j \neq 0$. This is seen to be a contradiction of (8) in view of §1.3 (ii). Thus the lemma is true.

THEOREM 4.13. *A homomorph $\bar{S}$ of a quasi-group S is a quasi-group.*

PROOF. Let H_i denote the complex composed of those elements of S which correspond to $\bar{a}_i$ in $\bar{S}$. Then by (i) $H_i \wedge H_j = 0$ if $i \neq j$. Now if in $\bar{S}$, $\bar{a}_i\bar{a}_j = \bar{a}_k$, then $H_iH_j \leqq H_k$ since every element a_ia_j of H_iH_j is mapped on $\bar{a}_k$ by (iii). By the lemma just proved we see that the H_i are all of the same order and that $H_iH_j = H_k$. We now show that $\bar{S}$ is a quasi-group. First, it satisfies §1.1 (i) by definition. Second, it satisfies §1.1 (ii) also since if $\bar{a}_i\bar{a}_j = \bar{a}_i\bar{a}_k$ in $\bar{S}$ for $j \neq k$ then in S we would have $H_iH_j = H_iH_k$ for $H_j \wedge H_k = 0$, which contradicts §1.3 (v).

We have already remarked that the quotient quasi-group S/H is a homomorph of S. The correspondence is $b \to Ha_i$ if $b < Ha_i$ and "product" in S/H is the complex product $(Ha_i)(Ha_j)$. We now show that the converse is true.

THEOREM 4.14. *Corresponding to any homomorph $\bar{S}$ of a quasi-group S there exists in S an invariant complex H such that S/H and $\bar{S}$ are isomorphic.*

PROOF. We saw in Theorem 4.13 that the elements of S fall into disjoint classes H_i of equal orders such that the correspondence $H_i \leftrightarrow \bar{a}_i$ is one-to-one and the H_i form a quasi-group isomorphic with $\bar{S}$. Thus $H_1H_1, H_1H_2, \cdots, H_1H_t$ are all distinct (§1.1 ii) and equal to $H_{i_1}, H_{i_2}, \cdots, H_{i_t}$ where the subscripts $i_1, i_2, \cdots, i_t$ are a permutation of $1, 2, \cdots, t$. Now the relation $H_{i_j} = H_1H_j = H_1a_j$ (where a_j is any element of H_j) shows that the H_i are all of the form H_1a_i. Then obviously $(H_1a)(H_1b) = H_1c$ for some c and H_1 is invariant by §4.1.

4.3. *Invariant complexes with special properties.* We have mentioned the existence (§4.1) of invariant complexes H such that

$$(Hx)(Hy) = H(xy) \tag{11}$$

for all x, y. We now ask: What is special about the homomorph $\bar{S}$ which corresponds to an invariant complex H satisfying (11)? The answer is found in the following

THEOREM 4.15. *A necessary and sufficient condition that there exist an invariant complex H in S satisfying* (11) *is that there exist a homomorph $\bar{S}$ of S containing an element $\bar{a}$ such that*

$$(\bar{a}\bar{a}_i)(\bar{a}\bar{a}_j) = \bar{a}(\bar{a}_i\bar{a}_j) \tag{12}$$

for all i, j and such that $a \to \bar{a}$ if $a < H$.

PROOF. The condition is necessary since S/H is a homomorph for which, if $x < Ha$, $y < Hb$, the relation $(Hx)(Hy) = H(xy)$ in S implies $(H(Ha))(H(Hb)) = H((Ha)(Hb))$ in S/H. Thus in $\bar{S}(= S/H)$ we have $\bar{a}(= H)$ with the property required by the theorem.

The condition is sufficient for if H, H_i, H_j denote the classes of elements in S which correspond to $\bar{a}$, $\bar{a}_i$, $\bar{a}_j$, respectively then (Theorem 4.13)

$$(HH_i)(HH_j) = H(H_iH_j) \tag{13}$$

for all i, j. Let now x, y be any elements of S. If $x < H_i$, $y < H_j$, then $xy < H_iH_j$ so that by §1.3 (iv) we have $(Hx)(Hy) = H(xy)$.

COROLLARY 1. *A necessary condition for* (12) *is that* $\bar{a}$ *be idempotent in* $\bar{S}$, *that is, that* $\bar{a}^2 = \bar{a}$.

For (12) implies $(\bar{a}\bar{a})(\bar{a}\bar{a}) = \bar{a}(\bar{a}\bar{a})$ so that (§1.1 ii) $\bar{a}^2 = \bar{a}$.

COROLLARY 2. *A sufficient condition for* (12) *is that* $\bar{a}$ *be a left identity in* $\bar{S}$.

For if $\bar{a}\bar{a}_i = \bar{a}_i$ for all i, then $(\bar{a}\bar{a}_i)(\bar{a}\bar{a}_j) = \bar{a}_i\bar{a}_j = \bar{a}(\bar{a}_i\bar{a}_j)$.

As an immediate consequence of Corollary 1, we have

THEOREM 4.16. *If an invariant complex* H *satisfies the relation* $(Hx)(Hy) = H(xy)$ *for all* x, y *in* S, *then* H *is a sub-quasi-group, i.e.,* $HH = H$.

And from Corollary 2 follows

THEOREM 4.17. *An invariant complex* H *such that*

$$H(Ha) = Ha \tag{14}$$

for all a *in* S *also satisfies the relation* (11).

We shall require the following theorem later.

THEOREM 4.18. *If* H *be an invariant complex such that* $H(Ha) = Ha$ *for all* a, *then* $a < Ha$ *and conversely; and either of these conditions implies* $(Ha)b = H(ab)$ *for all* a, b *in* S.

PROOF. Let $H(Ha) = Ha$. Suppose $a < Hb$; then by §1.3 (iv), $Ha = H(Hb)$ so that by Theorem 4.1, $Hb = Ha$. Conversely, if $a < Ha$ then $Ha = H(Ha)$. Finally, if $b < Hb$ then $(Ha)(Hb) = (Ha)b = H(ab)$ by the first part of this theorem and by Theorem 4.17.

COROLLARY. $H(Kb) = (HK)b$ *for any complex* K.

We now consider briefly the relation

$$(xH)(yH) = (xy)H. \tag{15}$$

Theorems analogous to Theorems 4.15–4.18 may be proved for this case if the order of all products Hx used in them be reversed. The following theorem relates the two types (11) and (15) of invariant complexes.

THEOREM 4.19. *If* H *be invariant in* S *and if*

$$H(Ha) = Ha = (Ha)H$$

for all a, *then* H *is normal in* S (§3.2) *and thus satisfies both* (11) *and* (15).

For we have $a < Ha$ by Theorem 4.18. Then $(Ha)H = aH$ by §1.3 (iv), so that by our hypothesis $aH = Ha$ for all a in S. Thus H is normal in S. The rest of the theorem is obvious.

4.4. *Relations between invariant complexes.*

THEOREM 4.20. *Let* K *be an invariant complex such that* $K > H$. *If* H *satisfies* (11) *so does* K *or if* H *satisfies* (14) *so does* K.

PROOF. Since K is invariant $(Ka)(Kb) = Kc$ for some c. Now $Ha < Ka$ and $Hb < Kb$. Thus $H(ab) = (Ha)(Hb) < Kc$. We have also $H(ab) < K(ab)$. Thus $Kc \wedge K(ab) \neq 0$. Therefore by Theorem 4.4, $Kc = K(ab)$. This

proves the first part of the theorem. The latter part follows from the fact that if $a < Ha$ and $H < K$, then $a < Ha < Ka$.

THEOREM 4.21. *If H and K both satisfy* (11) *then their intersection D, if not vacuous, also satisfies* (11). *And if H and K satisfy* (14) *then D is not vacuous and satisfies* (14).

PROOF. Since $(Da)(Db) < (Ha)(Hb)$ we have $(Da)(Db) < H(ab)$. Similarly $(Da)(Db) < K(ab)$. But since $D = H \wedge K$ we have from §1.3 (ii) $D(ab) = H(ab) \wedge K(ab)$. Thus $(Da)(Db) \leqq D(ab)$ with the inclusion sign impossible by §1.3 (i). Hence $(Da)(Db) = D(ab)$.

If $a < Ha$, $a < Ka$, then $a < (H \wedge K)a = Da$. Now $D = (H \wedge K) = 0$ is impossible since then $Ha \wedge Ka = 0$ which contradicts the fact that both contain a.

THEOREM 4.22. *If H satisfies* (14) *and K satisfies* (11), *then $HK = M$ satisfies* (11) *and $K < M$. Also if K satisfies* (14) *so does M.*

PROOF. We have $(Ma)(Mb) = ((HK)a)((HK)b) = (H(Ka))(H(Kb))$ by the corollary to Theorem 4.18. Now $(H(Ka))(H(Kb)) = H((Ka)(Kb))$ since H satisfies (11) also. But

$$H((Ka)(Kb)) = H(K(ab)) = (HK)(ab) = M(ab).$$

Thus M satisfies (11). Now since $k < Hk$ for all $k < K$ we have $K < HK = M$. Finally, if K satisfies (14), $a < Ka$ for any a. Thus from the preceding steps we have $a < H(Ka) = Ma$.

4.5. We now give a brief discussion of maximal invariant complexes and simple quasi-groups.

DEFINITION. An invariant complex is called *maximal* in S if it is contained in no other invariant complex except S itself.

DEFINITION. A quasi-group is called *simple* if it contains no invariant complexes (that is, none consisting of more than one element of S).

THEOREM 4.23. *If H be a maximal invariant complex then S/H is simple.*

PROOF. Let S/H contain an invariant complex. If this complex does not contain H, then one of its co-sets does. But by Theorem 4.8 this co-set is an invariant complex. Hence H is not maximal in S by Theorem 4.12. This contradicts our hypothesis.

THEOREM 4.24. *If H and K are permutable* (*that is, $HK = KH$*), *if both satisfy* (14), *and if both are maximal, then if we denote HK by M and $H \wedge K$ by D we have $M = S$ and S/H isomorphic with K/D, S/K isomorphic with H/D.*

PROOF. By Theorem 4.22, M is invariant and contains H and K since they are permutable. Now this means that $M = S$ since H and K are maximal. By Theorem 4.21, $D \neq 0$ and satisfies (11). Since $D < K$ and $(Dk_1)(Dk_2) = D(k_1k_2)$ we see that D is invariant in K (that K is a quasi-group follows from Theorem 4.16). Thus as in (6) we have

$$K = Dk_1 + Dk_2 + \cdots + Dk_t. \tag{16}$$

Then

$$HK = H(Dk_1) + H(Dk_2) + \cdots + H(Dk_t) \tag{17}$$
$$= (HD)k_1 + (HD)k_2 + \cdots + (HD)k_t$$

by the corollary to Theorem 4.18. Now $HD = H$ since $HD \leqq H$ (for $D < H$) with the inclusion sign impossible by §1.3 (i). Now from (17), since we have already proved $HK = S$, we have

$$S = HK = Hk_1 + Hk_2 + \cdots + Hk_t. \tag{18}$$

We compare (18) with (16) and find that the correspondence $Hk_i \leftrightarrow Dk_i$ is an isomorphism between S/H and K/D since both H and D satisfy (11).

The other part of the theorem follows similarly.

College of the City of New York, The City College.

Annals of Mathematics
Vol. 41, No. 3, July, 1940

ÜBER EINE ALLGEMEINE THEORIE DER FUCHSSCHEN GRUPPEN UND THETA-REIHEN

Von Masao Sugawara

(Received June 6, 1939)

Die Theorie der elliptischen Modulfunktionen ist kürzlich von Herrn C. L. Siegel auf den Fall verallgemeinert worden, der sich auf allgemeinen Typus der Riemannschen Flächen mit dem Geschlecht n bezieht.[1] Entsprechend möchte ich hier die Theorie der Poincaréschen automorphen Funktionen[2] auf "höherdimensionale" Fälle übertragen. Wir führen den Begriff der allgemeinen *Fuchsschen Gruppen* ein, und beweisen dafür den *Diskontinuitätssatz,* auf Grund dessen die Konvergenz der für diese Gruppen konstruierten *Theta-Reihen* gesichert sein wird. Die Siegelsche Modulgruppe steht in enger Beziehung mit einer gewissen Fuchsschen Gruppe. Insbesondere folgt die Existenz des Fundamentalbereiches für die Siegelsche Gruppe direkt aus unserem Diskontinuitätssatz.[3]

I. Bezeichnungen und Vorbemerkungen

Es handelt sich im folgenden hauptsächlich um n-dimensionale quadratische Matrizen mit komplexen Elementen. Mit grossen lateinischen Buchstaben bezeichnen wir meistenfalls solche Matrizen. Schreiben wir etwa

$$U = \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix}, \text{ bzw. } \quad X = \begin{pmatrix} P \\ Q \end{pmatrix},$$

so bedeuten dabei U_1, U_2, U_3, U_4, P, Q immer n-dimensionale Matrizen; U ist die daraus gebildete $2n$-dimensionale Matrix, und X eine Matrix vom Typus $(2n, n)$. Mit E bezeichnen wir die (n-dimensionale) Einheitsmatrix. Die transponierte bzw. konjugiert-komplexe Matrix von einer Matrix A sei mit A' bzw. $\bar{A}$ bezeichnet.

Ist H Hermitesch, und ist die zugehörigen Hermitesche Form positiv definit, so schreiben wir $H > 0$. Sind H_1, H_2 Hermitesch und $H_1 - H_2 > 0$, schreiben wir auch $H_1 > H_2$. Da offenbar mit $H_1 > 0$, $H_2 > 0$ auch $H_1 + H_2 > 0$ gilt, kann man mit diesen "Ungleichungen" wie üblich rechnen. Ist A regulär, d. h.

[1] C. L. Siegel: *Analytische Theorie der quadratischen Formen.* Ann. of Math. 36, Insb. §13.

[2] H. Poincaré: Acta Math. I, II, III. Œuvre, t. II. Vgl. auch. H. Weyl: *Die Idee der Riemannschen Fläche,* insb. §§20-21.

[3] Bei der Redaktion dieser Arbeit ist mir Herr S. Iyanaga behilflich gewesen. Ich möchte ihm auch an dieser Stelle meinen besten Dank aussprechen.

ist die Determinante $|A| \neq 0$, so gilt stets $A'\bar{A} > 0$. Ist umgekehrt $H > 0$, so lässt sich H in der Form $H = A'\bar{A}$ mit einem regulären A darstellen.

Ein Paar von Matrizen P, Q heisst nach Siegel[4] symmetrisch, falls $P'Q = Q'P$ gilt. Ist dabei Q regulär, so ist die Matrix PQ^{-1} symmetrisch. Umgekehrt lässt sich jede symmetrische Matrix Z in der Form. $Z = PQ^{-1}$, $|Q| \neq 0$ schreiben, wobei P, Q ein symmetrisches Matrizenpaar bilden: es genügt ja etwa $P = Z$, $Q = E$ zu setzen.

II. Raum, Grundgebiet und Bewegungen

Die Gesamtheit aller symmetrischen Matrizen Z bildet unsern *Raum* $\mathfrak{R}$. Jedes Z heisst ein Punkt von $\mathfrak{R}$. Wir wollen uns dabei $\mathfrak{R}$ als im $n(n+1)$-dimensionalen Euklidischen Raum eingebettet denken: $\mathfrak{R}$ ist also metrisch und als Entfernung $\rho(Z, Z^0)$ zweier Punkte $Z = (z_{ik}) = (x_{ik} + iy_{ik})$ und $Z^0 = (z^0_{ik}) = (x^0_{ik} + iy^0_{ik})$ gilt die Grösse $[\sum_{i \geq k} ((x_{ik} - x^0_{ik})^2 + (y_{ik} - y^0_{ik})^2)]^{\frac{1}{2}}$. Die Menge $\mathfrak{A}$ der Punkte $Z = (z_{ik})$, so dass $E > Z'\bar{Z}$ gilt, nennen wir das *Grundgebiet*. $\mathfrak{A}$ ist ersichtlich eine offene Menge in $\mathfrak{R}$. $\mathfrak{A}$ ist ferner beschränkt.

Aus $E > Z'\bar{Z}$ folgt nämlich, dass die Elemente auf der Hauptdiagonale von $E - Z'\bar{Z}$ positiv sein müssen, d. h. $\sum_{k=1}^{n} |z_{ik}|^2 < 1$, $i = 1, 2, \cdots, n$.[5]

Nun setzen wir

$$J = \begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix}.$$

Die $2n$-dimensionalen (regulären) Matrizen U, die den Bedingungen

$$(1) \qquad U'JU = J, \qquad\qquad (2) \qquad U'S\bar{U} = S$$

genügen, bilden eine Gruppe Γ. Sei U ein Element von Γ. Setzen wir $U = \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix}$ und betrachten wir die Substitution

$$(3) \qquad Z_1 = (U_1 Z + U_2)(U_3 Z + U_4)^{-1}.$$

Satz 1. *Die Substitution* (3) *ist für ale* $Z \in \mathfrak{A}$ *sinnvoll und bildet* $\mathfrak{A}$ *in sich ab.*

Demnach nennen wir die Substitution (3) eine *Bewegung* von $\mathfrak{A}$, die von $U \in \Gamma$ induziert wird.

Zum Beweis schreiben wir $Z = PQ^{-1}$ mit ein symmetrisches Paar bildenden Matrizen P, Q, $|Q| \neq 0$ und setzen $X = \begin{pmatrix} P \\ Q \end{pmatrix}$, $X_1 = UX$, $X_1 = \begin{pmatrix} P_1 \\ Q_1 \end{pmatrix}$. Nach (1), gilt

$$X_1'JX_1 = X'U'JUX = X'JX,$$

[4] Siegel a. a. O. §12.

[5] Übrigens ist $\mathfrak{A}$ auch konvex. Aus $E > Z_1'\bar{Z}_1$, und $E > Z_2'\bar{Z}_2$ folgt nämlich $E > \left(\frac{Z_1 + Z_2}{2}\right)'\overline{\left(\frac{Z_1 + Z_2}{2}\right)}$ wegen

$$4\left\{E - \left(\frac{Z_1 + Z_2}{2}\right)'\overline{\left(\frac{Z_1 + Z_2}{2}\right)}\right\} = 2(E - Z_1'\bar{Z}_1) + 2(E - Z_2'\bar{Z}_2) + (Z_1 - Z_2)'\overline{(Z_1 - Z_2)}.$$

oder
$$P_1'Q_1 - Q_1'P_1 = P'Q - Q'P = 0.$$

D. h. P_1, Q_1 bilden ein symmetrisches Matrizenpaar. Aus (2) folgt ähnlicherweise
$$X_1'S\bar{X}_1 = X'U'S\bar{U}\bar{X} = X'S\bar{X},$$
oder
$$(*) \quad P_1'\bar{P}_1 - Q_1'\bar{Q}_1 = P'\bar{P} - Q'\bar{Q}.$$

Da nun $E > Z'\bar{Z}$ mit $Q'\bar{Q} > P'\bar{P}$ gleichbedeutend ist, folgt $Q_1'\bar{Q}_1 > P_1'\bar{P}_1$. Ist andererseits $Q_1\varphi = 0$ mit n-dimensionalem Vektor φ, so auch $\varphi'Q_1' = 0$ und (*) liefert durch linksseitige Multiplikation mit φ' und rechtsseitige Multiplikation mit $\bar{\varphi}$ die Gleichung
$$(P_1\varphi)'(\overline{P_1\varphi}) = \varphi'(P'\bar{P} - Q'\bar{Q})\bar{\varphi};$$
die linke Seite dieser Gleichung ist $\geqq 0$, während die rechte Seite $\leqq 0$ ist; also $\varphi'(P'\bar{P} - Q'\bar{Q})\bar{\varphi} = 0$. Mithin $\varphi = 0$, weil $P'\bar{P} - Q'\bar{Q} < 0$, daher ist $| Q_1 | \neq 0$.

Die Substitution (3) ist also sinnvoll für $Z \in \mathfrak{A}$ und $Z_1 = P_1Q_1^{-1}$. Hieraus folgt dass mit $Z \in \mathfrak{A}$ auch $Z_1 \in \mathfrak{A}$ gilt.

Satz 2. *Jedes $U \in \Gamma$ hat folgende Gestalt.*
$$(4) \qquad U = \begin{pmatrix} U_1 & U_2 \\ \bar{U}_2 & \bar{U}_1 \end{pmatrix},$$
wobei gilt:

(5) $U_1'\bar{U}_2 = \bar{U}_2'U_1$ *d. h. U_1 und $\bar{U}_2$ bilden ein symmetrisches Matrizenpaar*; *und*
$$(6) \qquad U_1'\bar{U}_1 - \bar{U}_2'U_2 = E.$$

Umgekehrt gehört jede solche Matrix der Gruppe Γ an.

Beweis. Eliminiert man U' aus (1), (2), so bekommt man $-JU^{-1}J = S\bar{U}^{-1}S$ oder $\bar{U} = -SJUJS$, oder auch, ausgeschrieben
$$\begin{pmatrix} \bar{U}_1 & \bar{U}_2 \\ \bar{U}_3 & \bar{U}_4 \end{pmatrix} = \begin{pmatrix} U_4 & U_3 \\ U_2 & U_1 \end{pmatrix}.$$
Daraus folgt, dass U die Form (4) hat. Setzt man das in (1) oder in (2) ein und rechnet man aus, erhält man die Bedingungen (5), (6). Umgekehrt bestätigt man leicht, dass eine Matrix U von der Form (4) mit den (5), (6) genügenden U_1, U_2 die Bedingungen (1), (2) erfüllt.

Die Bewegung (3) schreibt sich hiernach in der Form
$$(3') \qquad Z_1 = (U_1Z + U_2)(\bar{U}_2Z + \bar{U}_1)^{-1},$$
wo U_1, U_2 (5), (6) genügen. Soll insbesondere eine Bewegung den Nullpunkt (d. h. die Nullmatrix) O festlassen, so muss in der Darstellung (3') $U_1'\bar{U}_1 = E$ (d. h. U_1 unitär), und $U_2 = 0$ sein. Setzt man nämlich $Z = Z_1 = 0$ in der aus (3') folgenden Formel $Z_1(\bar{U}_2Z + \bar{U}_1) = U_1Z + U_2$, erhält man $U_2 = 0$; $U_1'\bar{U}_1 = E$ folgt dann aus (6). Umgekehrt eine solche Substitution ersichtlich genügt die Bedingung.

Lässt eine Bewegung sogar alle Punkte von $\mathfrak{A}$ fest, d. h. ist sie die identische Bewegung von $\mathfrak{A}$, so muss ferner U_1 eine diagonale Matrix mit den Diagonalelementen ± 1 sein. Dann gilt nämlich $Z = U_1 Z \bar{U}_1^{-1}$ oder $Z\bar{U}_1 = U_1 Z$ für alle $Z \in \mathfrak{A}$. Also ist U_1 reell und mit allen $Z \in \mathfrak{A}$ vertauschbar, woraus leicht die Behauptung folgt.—Die aus solchen U_1 und $U_2 = 0$ gebildeten U bilden einen Normalteiler Γ_0 mit der Ordnung 2^n von Γ. Die Gruppe $\mathfrak{B}$ aller Bewegungen ist $\cong \Gamma/\Gamma_0$.

Sei nun A ein Punkt von $\mathfrak{A}$. Man kann eine Bewegung, die A in O überführt, folgendermassen konstruieren: Man setze $E - A'\bar{A} = C'\bar{C}$, $(C')^{-1} = C_1$ und

$$U_A = \begin{pmatrix} C_1 & 0 \\ 0 & \bar{C}_1 \end{pmatrix} \begin{pmatrix} E & -A \\ -\bar{A} & E \end{pmatrix} = \begin{pmatrix} C_1 & -C_1 A \\ -\bar{C}_1 \bar{A} & \bar{C}_1 \end{pmatrix}.$$

Dieses U_A gehört zu Γ: es hat ja die Form (4) und die Bedingungen (5), (6) sind für $U_1 = C_1$, $U_2 = -C_1 A$ erfüllt, wie man rechnerisch verifizieren kann.[6]

Die entsprechende Bewegung

$$Z_1 = C_1(Z - A)(E - \bar{A}Z)^{-1}\bar{C}_1^{-1}$$

führt offenbar A in O über. Hieraus folgt der

Satz 3. *Die allgemeine Bewegung von $\mathfrak{A}$, die einen Punkt $A \in \mathfrak{A}$ in einen Punkt $B \in \mathfrak{A}$ überführt, wird durch das Element* $U_B^{-1}\begin{pmatrix} U_1 & 0 \\ 0 & \bar{U}_1 \end{pmatrix} U_A$ *von* Γ, $(U_1'\bar{U}_1 = E)$, *induziert. Sie wird also durch folgende Formel gegeben:*

$$(7) \qquad D_1(Z_1 - B)(E - \bar{B}Z_1)^{-1}\bar{D}_1^{-1} = U_1 C_1 (Z - A)(E - \bar{A}Z)^{-1}\bar{C}_1^{-1} U_1',$$

wobei C_1 die frühere Bedeutung, und D_1 die entsprechende Bedeutung für B hat. Lässt man hier U_1 alle unitäre Matrizen durchlaufen, bekommt man alle solche Bewegungen.[7]

Die Bewegung (7) wollen wir kurz mit (A, B, U_1) bezeichnen.

III. Allgemeine Fuchssche Gruppe

Die Bewegungsgruppe $\mathfrak{B}$ ist eine topologische Gruppe: die Bewegungen $(A_1, B_1, U_1^{(1)})$ und $(A_2, B_2, U_1^{(2)})$ liegen nahe bei einander, wenn sich die Matrizen $A_1, B_1, U_1^{(1)}$ bzw. von $A_2, B_2, U_1^{(2)}$ nur um wenig unterscheiden. Eine Untergruppe $\mathfrak{G}$ von $\mathfrak{B}$ nennen wir eine *allgemeine Fuchssche Gruppe*, wenn sie keine infinitesimale Bewegung enthält, d. h. wenn das Einselement (mithin auch jedes Element) in der Gruppenmannigfaltigkeit isoliert liegt.

[6] (5) besagt: $C_1'\bar{C}_1\bar{A} = \bar{A}\bar{C}_1'C_1$, oder $\bar{A}C'\bar{C} = \bar{C}'C\bar{A}$. Das folgt aus $\bar{A}(E - A\bar{A}) = (E - \bar{A}A)\bar{A}$, indem man $A' = A$ beachtet. (6) besagt: $C_1'\bar{C}_1 - \bar{A}\bar{C}_1'C_1A = E$ oder $(\bar{C}'C)^{-1} - \bar{A}(C'\bar{C})^{-1}A = E$. Das folgt aus $E - \bar{A}A = E - \bar{A}(E - A\bar{A})^{-1}A(E - \bar{A}A)$.

[7] Setzt man $Z = A + dA$ in (7), und lässt man dA nach O rücken, findet man, dass $ds^2 = \mathrm{Sp}\, dA(E - \bar{A}A)^{-1}\, d\bar{A}'(E - A\bar{A})^{-1}$ bewegungsinvariant ist (indem man die Glieder mit $(dA)^2$ vernachlässigt). So kann man $\mathfrak{A}$ als einen Riemannschen Raum mit diesem ds^2 auffassen. In dieser Note machen wir aber davon noch keinen Gebrauch. Vgl. R. Nevanlinna: *Eindeutige analytische Funktionen* §1.

Satz 4. (*Diskontinuitätssatz.*) *Jede allgemeine Fuchssche Gruppe ist eigentlich diskontinuierlich in $\mathfrak{A}$. Ist nämlich A ein beliebiger Punkt von $\mathfrak{A}$, und sind σ_i die Elemente von einer allgemeinen Fuchsschen Gruppe $\mathfrak{G}$, so weist die Punktmenge $\{\sigma_i A\}$ keinen Häufungspunkt in $\mathfrak{A}$ auf.*

Beweis.[8] Angenommen, die Punktmenge $\{\sigma_i A\}$ besitze eine Häufungsstelle *in* $\mathfrak{A}$. Setzen wir einfachheitshalber $\sigma_i A = A_i$. Für ein beliebiges ϵ gibt es dann ein gewisses i, sodass

$$\rho(A_i, A_k) < \epsilon \tag{8}$$

für unendlich viele Werte von k gilt. Die Bewegung $\sigma_k \sigma_i^{-1} \epsilon \mathfrak{G}$ führt den Punkt A_i in A_k über; wir können also $\sigma_k \sigma_i^{-1} = (A_i, A_k, U_1^{(k)})$ mit einem gewissen unitären $U_1^{(k)}$ schreiben. Da aber die Gruppe aller n-dimensionalen unitären Matrizen offenbar kompakt ist, lässt sich eine Teilfolge von $U_1^{(k)}$ so wählen, dass sie konvergiert. Wählt man also ein geeignetes Paar l, m der Werte von k aus dieser Teilfolge, so unterscheiden sich die Matrizen $U_1^{(l)}$ und $U_1^{(m)}$ beliebig wenig von einander. Wegen (8) liegen aber auch A_l, A_m beliebig nahe bei einander. Die Bewegung

$$\sigma_l \sigma_m^{-1} = (\sigma_l \sigma_i^{-1})(\sigma_m \sigma_i^{-1})^{-1} = (A_i, A_l, U_1^{(l)})(A_i, A_m, U_1^{(m)})^{-1} \epsilon \mathfrak{G}$$

wäre dann also infinitesimal, was unserer Voraussetzung widerspricht.

IV. Theta-Reihen

Es sei $\mathfrak{G}$ eine allgemeine Fuchssche Gruppe. Ein Element $\sigma \epsilon \mathfrak{G}$ sei von der Form:

$$\sigma(z) = (U_1 Z + U_2)(U_3 Z + U_4)^{-1}.$$

Die Theta-Reihen für $\mathfrak{G}$ definieren wir dann durch

$$\Theta_k(Z) = \sum_{\sigma \epsilon \mathfrak{G}} | U_3 Z + U_4 |^{-k(n+1)}, \qquad (k \geqq 2).$$

Satz 5. *Die Reihe $\Theta_k(Z)$ ist absolut und gleichmässig konvergent in der Umgebung von $Z \epsilon \mathfrak{A}$. Sie stellt also eine in $\mathfrak{A}$ reguläre analytische Funktion von z_{ik} dar, und es gilt:*

$$\Theta_k(\sigma(z)) = | U_3 Z + U_4 |^{k(n+1)} \Theta_k(Z).$$

Beweis. Es genügt offenbar die Konvergenz der Reihe zu zeigen. Es sei $Z^0 = (z_{ik}^0) \epsilon \mathfrak{A}$ und $\mathfrak{B} = (Z; | z_{ik} - z_{ik}^0 | < r)$.[9]
$r > 0$ sei dabei so klein gewählt, dass alle Gebiete $\sigma\mathfrak{B}$, $\sigma \epsilon \mathfrak{G}$, zu einander fremd liegen. $v(\mathfrak{B})$ sei das Euklidische Mass von $\mathfrak{B}$, welches offenbar positiv ist. Wegen der Beschränktheit von $\mathfrak{A}$ ist dann $\sum_{\sigma \epsilon \mathfrak{G}} v(\sigma\mathfrak{B})$ konvergent. Andererseits gilt $v(\sigma\mathfrak{B}) = \int_{\mathfrak{B}} I\, dZ$, wo $I = \left| \dfrac{\partial \sigma(Z)}{\partial Z} \right|$ den absoluten Betrag der Funk-

---

[8] Vgl. E. Cartan: *Leçons sur la géométrie projective complexe*, §78. Den Hinweis auf diesem Buch verdanke ich Herrn Iyanaga.

[9] $\mathfrak{B}(Z; | z_{ik} - z_{ik}^0 | < r)$ bedeutet die Menge derjenigen Z, für welche $| z_{ik} - z_{ik}^0 | < r$.

tionaldeterminate für die Bewegung σ bedeutet. Man setze nun $P_1 = U_1Z + U_2$, $Q_1 = U_3Z + U_4$ sodass $Z_1 = P_1Q_1^{-1}$, $P_1 = Z_1Q_1$, also $dP_1 = dZ_1Q_1 + Z_1\,dQ_1$. Daher folgt $Q_1'\,dZQ_1 = Q_1'\,dP_1 - Q_1'Z_1\,dQ_1 = Q_1'\,dP_1 - P_1'\,dQ_1 = Q_1'U_1\,dZ - P_1'U_3\,dZ = dZ$, letzteres wegen (4), (5), (6) Satz 3. Man hat also $dZ_1 = Q_1'^{-1}\,dZQ_1^{-1}$. Daraus findet man wie Siegel als Werte von I, $I = \| U_3Z + U_4 \|^{-2(n+1)}$.[10]

Hieraus folgt

$$v(\sigma\mathfrak{B}) \geqq \left| \int_{\mathfrak{B}} \frac{dP}{| U_3Z + U_4 |^{2(n+1)}} \right| = \frac{(\pi r^2)^{\frac{1}{2}n(n+1)}}{\| U_3Z_0 + U_4 \|^{2(n+1)}},\text{[11]}$$

$$\| U_3Z_0 + U_4 \|^{-2(n+1)} \leqq v(\sigma\mathfrak{B})(\pi r^2)^{-\frac{1}{2}n(n+1)};$$

also folgt die Konvergenz der Reihe $\Theta_2(Z_0)$. Daraus folgert man leicht die absolute und gleichmässige Konvergenz von $\Theta_2(Z)$ für $| z_{ik} - z_{ik}^0 | \leqq \frac{1}{2}r$. Man setze nun $S = \sum \| U_3Z + U_4 \|^{-2(n+1)}$. Dann gilt $\| U_3Z + U_4 \|^{-(n+1)} < S^{\frac{1}{2}}$, mithin $\| U_3Z + U_4 \|^{-k(n+1)} < \| U_3Z + U_4 \|^{-2(n+1)} S^{\frac{1}{2}(k-2)}$ für $k > 2$. Also konvergiert $\Theta_k(Z)$ auch für $k > 2$ absolut und gleichmässig.

V. Modulgruppe

Die von Siegel eingeführte Modulgruppe $\mathfrak{M}$ ist die Gruppe aller regulären $2n$-dimensionalen Matrizen mit ganz-rationalen Elementen, die der Bedingung (1) genügen. Sei $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathfrak{M}$. Ist $Z = (z_{ik}) = (x_{ik} + iy_{ik}) \in \mathfrak{R}$ und setzt man $X = (x_{ik})$, $Y = (y_{ik})$, sodass $Z = X + iY$, dann führt die Substitution $Z_1 = (AZ + B)(CZ + D)^{-1}$ ein Z mit $Y > 0$ in ein ebensolches $Z_1 = X_1 + iY_1$ mit $Y_1 > 0$ über. Die Gesamtheit $\mathfrak{H}$ aller $Z = X + iY$ mit $Y > 0$ nennen wir die *Siegelsche Halbebene*. $\mathfrak{H}$ bleibt also gegen Siegelschen Modulsubstitutionen invariant.

Mittels der transformation

$$W = (Z - iE)(Z + iE)^{-1}$$

wird nun $\mathfrak{H}$ auf unser Grundgebiet $\mathfrak{A}$ homöomorph abgebildet, wie folgende leicht verifizierbare Identität zeigt:

$$W'\overline{W} - E = -2(Z + iE)^{-1}\,Y\,\overline{(Z + iE)}^{-1}.$$

Setzt man entsprechend $T = \frac{1}{[2i]^{\frac{1}{2}}} \begin{pmatrix} E & -iE \\ E & iE \end{pmatrix}$, $U = TMT^{-1}$ für $M \in \mathfrak{M}$, so

---

[10] $\| A \|$ bedeutet den absoluten Betrag der Determinante von A. C. L. Siegel: *Über die Zetafunktionen indefiniter quadratischer Formen II.* Math. Zeit. 44, S. 403.

[11] Man erhält dies nach wiederholter Anwendung der Formel für eine analytischen Funktion f einer Variablen, $f(\zeta)r^2 = \int_0^r \int_0^{2\pi} f(\zeta + re^{i\theta})\,dr\,d\theta$, die sich aus der Integraldarstellung von Cauchy ergibt.

folgt $U \in \Gamma$.[12] Da ferner $\mathfrak{M}$ offenbar keine infinitesimale Transformation enthält, entspricht der Gruppe der Siegelschen Substitutionen eine allgemeine Fuchssche Gruppe. Nach unserem Satz 4 besitzt sie also einen Fundamentalbereich in der Siegelschen Halbebene.

MEGUROKU KOMABA DAIICHI KOTO-GAKKO,
TOKYO, JAPAN.

[12] Aus $T'JT = J$ folgt $U'JU = T^{-1'}M'T'JTMT^{-1} = T^{-1'}M'JMT^{-1} = J$. Aus $T'S\bar{T} = iJ$ folgt $U'S\bar{U} = T^{-1'}M'T'S\bar{T}M\bar{T}^{-1} = iT^{-1'}M'TM\bar{T}^{-1} = S$.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

FINITELY ADDITIVE INTEGRAL

BY S. BOCHNER

(Received November 16, 1939)

H. Freudenthal[1] has shown that under proper assumptions a partially ordered vector space C can be realized by measurable (countably additive set) functions on a properly construed Boolean algebra with a Lebesgue measure. His assumptions are that any countable set in C which is bounded from above (below) has a l.u.b. (g.l.b.); that there is another limit in the space whose topological implications are practically those of a Banach norm; and that there exists a positive element $F = 1$ no part of which is 0; precisely, if $g > 0$ then inf $(g, 1) > 0$.[2]

Generalizing a previous construction[3] we will establish an analogous result for the *finitely* additive case. We will completely eliminate the assumption that infinite sets of element which are bounded in the sense of order have an inf or sup respectively. Instead of the limit topology we will more specifically assume the existence of a functional with the properties of an integral, but it will be only finitely additive. However our assumption concerning the element $f = 1$ will be rather restrictive; we will assume that every other element g is bounded from above and below by a numerical multiple of 1.

Our construction is given in section IV after a preparatory step in section III. Sections I and II contain definitions and obvious analogues to other results given in "A." Furthermore we have appended to each theorem a generalization bearing on the case in which the integral or measure involved is not a number but an element of another vector space. It is worth pointing out that the conditions which have to be imposed on that vector space differ seriously from case to case.

I. Definitions

A partially ordered space C has the following properties:

1) C *is an Abelian group with respect to addition and with real numbers as coefficients (operators).*

2) C *is partially ordered (relation* $>$*) and* (i) *corresponding to any elements* f, g *there exists an element* sup (f, g) *which is the smallest element* $\geqq f$, $\geqq g$, *and a similar element* inf (f, g), (ii) $f > 0$ *implies* $0 > -f$, $f > g$ *implies* $f + h > g + h$, (iii) $f > 0$, $a > 0$ *(a is a real number) implies* $af > 0$.

---

[1] H. Freudenthal, *Teilweise geordnete Moduln*, Proceedings Royal Academy Amsterdam, 36(1936), 641–651.

[2] We are using the notation of L. V. Kantorovitch, *Lineare halbgeordnete Raüme*, Recueil Mathématique, Moscow, 2(1937), 121–168.

[3] S. Bochner, *Additive set functions on groups*, Annals of Mathematics, 40(1939), 769–799, esp. 774–775. This paper which will be quoted frequently will be referred to as "A".

3) *There exists a positive element* 1 (*its multiple with a number a will be also denoted by a*) *such that corresponding to any* $f \in C$ *there exist numbers* p, q *for which*

$$p \leqq f \leqq q. \tag{1}$$

A *field* $[C, M]$ has the further property:

4) Mf *is a number which is defined for all* $f \in C$ *such that* (i) $M(af + bg) = aMf + bMg$, (ii) $Mf > 0$ *for* $f > 0$, (iii) $M1 = 1$.

A *complete* field $[C, M]$ has the additional property

5) *If* f_n, $g_n \in C$ *such that*

$$f_n \geqq f_{n+1}, \qquad g_{n+1} \geqq g_n, \qquad f_m \geqq g_n \qquad (m, n = 1, 2, \cdots) \tag{2}$$

$$\lim_{m \to \infty} Mf_m = \lim_{n \to \infty} Mg_n \tag{3}$$

then there exists one and only one element h, *such that*

$$f_m \geqq h \geqq g_n \qquad (m, n = 1, 2, \cdots). \tag{4}$$

The element h is unique. If h and h_1 both satisfy (4) then so will their sup and inf. But

$$f_m \geqq \sup (h, h_1) \geqq \inf (h, h_1) \geqq g_n$$

implies

$$M (\sup (h, h_1) - \inf (h, h_1)) \leqq Mf_m - Mg_n$$

and therefore, by (3) and 4), (ii),

$$\sup (h, h_1) = \inf (h, h_1).$$

Hence $h = h_1$.

A *Boolean algebra* B of elements $\{E\}$ has the property

6) B *is a partially ordered set with a smallest element* 0 *and a largest element* $E = B$, *and two operations of "addition" and "multiplication" which are commutative and associative each and distributive with each other and an operation of complementation such that* (i) $E_1 \leqq E_2$ *if and only if* $E_1 = E_1E_2$, (ii) $E + E = EE = E$, (iii) $E_1 = E_1(E_1 + E_2) = E_1 + E_1E_2$, (iv) *corresponding to elements* $E_1 \leqq E_2$ *there exists an element* $E_3 (= E_2 - E_1)$ *for which* $E_1E_3 = 0$, $E_1 + E_3 = E_2$.

A (*Jordan*) *field* $[B, vE]$ has the further property

7) $v0 = 0$, $vE > 0$ *if* $E > 0$, $vB = 1$, $v(E_1 + E_2) = vE_1 + vE_2$ if $E_1E_2 = 0$. *The measure* vE *will also be denoted by* $| E |$.

A *complete field* $[B, vE]$ has the additional property

8) *If* E'_n, $E''_n \in B$ *such that*

$$E'_n \geqq E'_{n+1}, \qquad E''_{n+1} \geqq E''_n, \qquad E'_m \geqq E''_n \qquad (m, n = 1, 2, \cdots) \tag{5}$$

$$\lim_{m \to \infty} | E'_m | = \lim_{n \to \infty} | E''_n | \tag{6}$$

then there exists one (and only one) element E such that

$$E'_m \geqq E \geqq E''_n. \tag{7}$$

II. Additive set functions on Jordan fields

The definitions and theorems given in "A," p. 773–783 can be made available for set functions on a general Jordan field $[B, vE]$ by two different procedures. We can either use the theorem of Stone[4] that any Boolean algebra can be represented by subsets of a point set; or we can avoid the transfinite induction involved in Stone's proof by duplicating all steps required. Since the second procedure requires minor modifications we will describe it briefly.

On any Boolean algebra B we again have the concept of (finite) partition $\delta = \delta(E) = (E_\nu)$, of refinement: $\delta < \delta'$ and of a monotone sequence of partitions $\{\delta_n\}$, see "A," p. 771. Every function $F(E)$ on B will be an additive set function of bounded variation the variation being defined as

$$\| F \|_1 = \lim_{\delta} \sum_{\nu=1}^{n} | F(E_\nu) |. \tag{8}$$

It is known[5] that the space of these functions satisfies our properties 1) and 2) if we define $F \geqq G$ to mean that $F(E) \geqq G(E)$ for all $E \in B$. In order to obtain other properties we have to consider a field $[B, vE]$. In this case we say that $F(E)$ is constant on E_0 if for all $0 < E < E_0$

$$\frac{F(E)}{|E|} = \frac{F(E_0)}{|E_0|}. \tag{9}$$

The common value of these quotients will again be called the value of F on E_0. A *step function* is one which has a constant value a_ν on each element E_ν of a suitable partition (E_ν). Given $A \in B$, the characteristic function $\omega_A(E)$ is the step function which has the value 1 on A and the value 0 on $B - A$. Accordingly a step function can be written in the form

$$F = \sum_{\nu=1}^{n} a_\nu \omega_{E_\nu}. \tag{10}$$

Defining now the element $F = 1$ to be the characteristic function ω_B and putting $MF = F(B)$ it is easily seen that the totality of step functions is a field $[C, M]$. There exists a smallest complete field $[R, M]$ containing $[C, M]$. An element F belongs to R if corresponding to any $\epsilon > 0$ there exist two step functions Φ, Ψ such that

$$\Phi \geqq F \geqq \Psi, \qquad M(\Phi - \Psi) < \epsilon. \tag{11}$$

---

[4] M. H. Stone, *The theory of representations for Boolean algebras*, Trans. Amer. Math. Soc., 40(1936), 36–111, esp. Theorem 67, p. 106.

[5] S. Saks, *Theory of the integral* (Stechart) 1937, p. 10 and p. 61.

The class R corresponds to the class of Riemann integrable functions.[6] The Darboux criterion for integrability is, assuming $\delta = (E_\nu)$,

$$\lim_{\delta} \sum_{\nu=1}^{n} \sup \left| \frac{F(E')}{|E'|} - \frac{F(E'')}{|E''|} \right| \cdot |E_\nu| = 0 \qquad (E', E'' \epsilon E_\nu). \tag{12}$$

A class R_0 which lies between C and R is defined by the severer requirement

$$\lim_{\delta} \max_{1 \leq \nu \leq n} \left| \frac{F(E')}{|E'|} - \frac{F(E'')}{|E''|} \right| = 0 \qquad (E', E'' \epsilon E_\nu). \tag{13}$$

However $[R_0, M]$ is not a complete field.

None of these fields is complete as a Banach space with the norm (8). A Banach space arises if we consider the space V_p, $1 \leq p < \infty$ of all functions with the norm

$$\| F \|_p = \left\{ \lim_{\delta} \sum_{\nu=1}^{n} \frac{|F(E_\nu)|^p}{|E_\nu|^{p-1}} \right\}^{1/p},$$

and the space V_∞ with the norm

$$\sup_E \frac{|F(E)|}{|E|}$$

and the subspace AC of V_1 consisting of functions which are absolutely continuous relative to the measure vE. In this connection, $F(E)$ is absolutely continuous if there exists an error function $\eta(\epsilon)$, $\lim_{\epsilon \to 0} \eta(\epsilon) = 0$, such that $|F(E)| \leq \eta(|E|)$.

If F is any function and $\delta = (E_\nu)$ is any partition, then the projection F_δ of F on δ is again the step function which on E_ν has the constant value $\dfrac{F(E_\nu)}{|E_\nu|}$. We again have the following theorem (see "A," theorems 5–10):

THEOREM 1. *In the Banach spaces AC and V_p, $p > 1$, the step functions are dense in norm. If $\{\delta_n\}$ is a monotone sequence of partitions then*

$$\lim \| F_{\delta_m} - F_{\delta_n} \| = 0 \qquad (m, n \to \infty).$$

In particular, if the space is separable it has a basis consisting of step functions.

The conjugate space to AC is V_∞. A sequence $\{F_\nu\}$ is weakly convergent in AC if the norms $\| F_\nu \|$ are bounded and the functions F_ν are uniformly absolutely continuous and the limit of $F_\nu(E)$ exists for each E.

The spaces V_p and V_q, $p > 1, q > 1, p = q/q - 1$ are conjugate. A sequence $\{F_\nu\}$, $F_\nu \epsilon V_p$ is weakly convergent if $\| F_\nu \| < K$ and $\lim F_\nu(E)$ exists for each E. Also if $\{F_\nu\}$ converges weakly to 0 then for a suitable subsequence F'_μ,

$$\left\| \sum_{\mu=1}^{m} F'_\mu \right\| = \begin{cases} 0(m^{1/p}) & \text{if } 1 < p \leq 2 \\ 0(m^{1/2}) & \text{if } p \geq 2. \end{cases} \tag{14}$$

[6] More precisely, this type of integral has been first defined in its general form by A. Kolmogoroff, *Untersuchungen über den Integralbegriff*, Math. Annalen, 103(1930), 654–696.

Finally we observe that the mapping process described in "A," pp. 780–781 can be applied to prove the following extension.[7]

THEOREM 1'. *Theorem 1 with the possible exception of relation* (14) *also holds if the values of our functions are not numbers but elements of a Banach space X which has the property that every function $f(t)$ from the interval $0 \leqq t \leqq 1$ to X which has bounded variation is differentiable almost everywhere.*

III. Completion of fields

THEOREM 2. *Corresponding to any field $[C, M]$ there exists a complete field $[\bar{C}, M]$ which is the smallest complete field containing it.*

We construct the space T of sequences

$$\varphi = \{f_k\} = \{f_1, f_2, \cdots\},$$

$\psi = \{g_k\}$, $\chi = \{h_k\}$ etc. with elements of C. We define: (i) $\varphi \geqq \psi$ if $f_k \geqq g_k$ for all $k \geqq k_0$ and (ii) $a\varphi + b\psi = \{af_k + bg_k\}$. This implies that $\varphi = \psi$ if and only if $f_k = g_k$ for all $k \geqq k_0$, that $0 = \{0\}$, and that (iii) $\sup(\varphi, \psi) = \{\sup(f_k, g_k)\}$, $\inf(\varphi, \psi) = \inf(f_k, g_k)$. The space T has the properties 1) and 2) and in addition the following property: If

$$\varphi_n \geqq \varphi_{n+1}, \qquad \psi_{n+1} \geqq \psi_n, \qquad \varphi_n \geqq \psi_n \qquad n = 1, \cdots$$

then there exists an element χ for which

$$\varphi_m \geqq \chi \geqq \psi_n.$$

In order to prove the latter property we set up sequences $\varphi_n = \{f_k^n\}$, $\psi_n = \{g_k^n\}$. Altering a finite number of elements f_k^n first for $n = 2$, then for $n = 3$, then for $n = 4$ etc., we may assume that $f_k^n \geqq f_k^{n+1}$ for all n and all k. Similarly we may assume that $g_k^{n+1} \geqq g_k^n$ for all n and all k. Now, corresponding to each n there exists an index $k(n)$ such that $f_k^n \geqq g_k^n$ for $k \geqq k(n)$; and it is easy to see that our property will hold if we put $\chi = \{h_k\}$ where

$$h_k = f_k^1 \quad \text{for} \quad k < k(2)$$

$$h_k = f_k^n \quad \text{for} \quad k(n) \leqq k < k(n+1).$$

The original space C will be a part of T if we identify any f of C with the sequence $\{f\}$.

Now let $\varphi \in T$. We temporarily call φ integrable if corresponding to any $\epsilon > 0$ there exist elements f, g of C such that

$$f \geqq \varphi \geqq g, \qquad M(f - g) < \epsilon. \tag{15}$$

The common value of inf Mf and sup Mg will be again denoted by $M\varphi$. We claim that the totality of integrable elements has all properties of a complete

[7] See S. Bochner and A. E. Taylor, *Linear functionals on certain spaces of abstractly-valued functions*, Annals of Math., 39(1938), 913–944; B. J. Pettis, *Differentiation in Banach spaces*, Duke Journal, 5(1939), 254–269.

field $[C_0, M]$ except for the one *deficiency* that $\varphi > 0$ implies $M\varphi \geqq 0$ and not necessarily $M\varphi > 0$. The reader will have no difficulty in proving the properties concerning $[C_0, M]$. As an illustration we will verify one of them, namely that if φ, φ_1 of T belong to C_0, then so does $\sup(\varphi, \varphi_1)$. In fact if $f \geqq \varphi \geqq g$ and $f_1 \geqq \varphi_1 \geqq g_1$ then

$$\sup(f, f_1) \geqq \sup(\varphi, \varphi_1) \geqq \sup(g, g_1)$$

and our assertion follows from

$$\begin{aligned} M(\sup(f, f_1) - \sup(g, g_1)) &= M \sup(f - \sup(g, g_1), f_1 - \sup(g, g_1)) \\ &\leqq M \sup(f - g, f_1 - g_1) \leqq M(f - g) + M(f_1 - g_1). \end{aligned}$$

In order to remove the deficiency mentioned before we apply a process of identification. Putting $2|\varphi| = \varphi + \sup(\varphi, 0)$, the totality of all elements $\varphi \in C_0$ for which $M(|\varphi|) = 0$ are a subspace S with properties 1) and 2) and we define $\bar{C}$ as the class of residues of C_0 mod S. Thus denoting for a while elements of C_0 by f, g etc., elements of S by $\varphi, \psi, \cdots$ etc. and elements of $\bar{C}$ corresponding to elements $f, g, \cdots$ by $\bar{f}, \bar{g}, \cdots$ respectively, then $f = ag + bh$, $f > g$, $f = \sup(g, h)$ shall mean that for appropriate elements φ, ψ, χ, $f = a(g + \varphi) + b(h + \psi)$, $f > g + \varphi$, $f + \chi = \sup(g + \varphi, h + \psi)$. Denoting the resulting field by $[\bar{C}, M]$ it is not hard to verify that it has all properties of a complete field.

The reader will easily establish the following extention of Theorem 2.

THEOREM 2′. *Theorem* 1 *also holds if the values of the functional* Mf *are not necessarily real numbers but elements of another partially ordered space which has properties* 1) *and* 2) *and in which every set of elements which is bounded from above (below) has a* sup (inf).

Another point worth mentioning is the following obvious

REMARK TO THEOREM 2. If C has properties 1)–3) and property 4) is satisfied for two different functionals M and M' and if the functionals M and M' are absolutely continuous each with respect to the other then the corresponding smallest completions C are the same.

The completion of Jordan fields could be proven directly in an analogous way.[8] However indirectly it is a consequence of theorem 2 and the following theorem 3.

IV. Generation of Jordan Fields

THEOREM 3. *Corresponding to any complete field* $[\bar{C}, M]$ *there exists a field* $[B, vE]$ *which is also complete such that* $\bar{C}$ *belongs to class* R_0 *on that field and* $Mf = f(B)$.

Also if $[C, M]$ *is a field of which* $[\bar{C}, M]$ *is the completion then the set* C *is dense in the spaces* AC *and* V_p, $p > 1$.

---

[8] H. M. MacNeille, *Extension of measure*, Proc. Nat. Acad. Science, 24(1938), 188–193.

PROOF. We consider all elements of C for which

$$0 \leqq \omega \leqq 1 \tag{16}$$

and

$$\inf(\omega, 1-\omega) = 0. \tag{17}$$

They give rise to a Boolean algebra B if we interpret sup as addition and inf as multiplication. The correspondence between an element of C and an element E of B will be expressed by the notations ω_E, E_ω. The measure $|E_\omega| = M\omega$ makes B into a field $[B, vE]$ whose characteristic functions can be readily identified with the elements ω_E of C we started from. In order to prove the completeness of our Jordan field we have to prove that the relations

$$\omega'_n \geqq \omega'_{n+1}, \qquad \omega''_{n+1} \geqq \omega''_n, \qquad \omega'_m \geqq \omega \geqq \omega''_m \qquad (m, n = 1, 2, \cdots)$$

$$\lim M\omega'_m = \lim M\omega''_m$$

$$\inf(\omega'_n, 1-\omega'_n) = 0, \qquad \inf(\omega''_n, 1-\omega''_n) = 0$$

imply (17). But this follows immediately from

$$M \inf(\omega'_n, 1-\omega) = M(\inf(\omega'_n, 1-\omega) - \inf(\omega'_n, 1-\omega'_n)) \leqq M(\omega'_n - \omega)$$

and

$$0 \leqq M \inf(\omega, 1-\omega) \leqq M \inf(\omega'_n, 1-\omega).$$

We next take a fixed element f of C and we introduce the elements

$$f(a) = \inf(f, a) \qquad -\infty < a < \infty \tag{18}$$

$$f(a, b) = \frac{f(b) - f(a)}{b-a} \qquad -\infty < a < b < \infty. \tag{19}$$

The function $f(a)$ is monotonely increasing (non-decreasing) with a, also

$$f(b) - f(a) = \inf(f, b) - \inf(f, a)$$

$$= \frac{f+b-|f-b|}{2} - \frac{f+a-|f-a|}{2} \leqq b-a$$

and thus

$$0 \leqq f(b, a) \leqq 1. \tag{20}$$

Furthermore, for $\alpha \geqq 0, \beta \geqq 0, \alpha + \beta = 1$,

$$f(\alpha a + \beta b) - \alpha f(a) - \beta f(b) = \{\inf(\alpha f, \alpha a + \beta b) - \inf(\alpha f, \alpha a)\}$$

$$+ \{\inf(\beta f, \alpha a + \beta b) - \inf(\beta f, \beta b)\} \geqq 0 + 0 = 0,$$

and thus $f(a)$ is concave. In particular we conclude that for arbitrary a, $h > 0$, $k > 0$, $f(a - h, a)$ is monotonely decreasing as h decreasing, $f(a, a + k)$ is monotonely increasing as k decreases, and

$$f(a - h, a) \geqq f(a, a + k). \tag{21}$$

Similar inequalities hold for the numerical functions

$$\varphi(a) = Mf(a), \qquad \varphi(a, b) = Mf(a, b).$$

In particular, since $\varphi(a)$ is concave its derivate exists at all but a countable number of exceptional values a. Thus, for non-exceptional values a

$$\lim_{h \to 0} \varphi(a - h, a) = \lim_{k \to 0} \varphi(a, a + k).$$

By property 5) this implies the existence of a unique element g_a for which

$$f(a - h, a) \geqq g_a \geqq f(a, a + k). \tag{22}$$

We are going to show that it is a characteristic function. Equation (16) follows immediately from (20) and (22). Furthermore we have, for $h > 0$,

$$\begin{aligned} h \cdot \inf\,&(f(a, a + h), 1 - f(a - h, a)) \\ &= \inf\,\{\inf\,(f, a + h) - \inf\,(f, a), h - \inf\,(f, a) + \inf\,(f, a - h)\} \\ &= -\inf\,(f, a) + \inf\,\{\inf\,(f, a + h), \inf\,(f + h, a)\} \\ &= -\inf\,(f, a) + \inf\,(f, a + h, f + h, a) \\ &= -\inf\,(f, a) + \inf\,(f, a) = 0, \end{aligned}$$

and therefore

$$\begin{aligned} 0 \leqq \inf\,(g_a, 1 - g_a) &\leqq \inf\,(f(a - h, a), 1 - f(a + h, a)) \\ &= \inf\,(f(a - h, a), 1 - f(a + h, a)) - \inf\,(f(a, a + h), 1 - f(a - h, a)) \\ &\leqq 2 \inf\,(f(a - h, a) - f(a, a + h)); \end{aligned}$$

these relations remain valid if we form the functional M for all terms. Doing so and letting $h \to 0$, we obtain $M \inf\,(g_a, 1 - g_a) = 0$ and thus g_a also satisfies relation (17).

The element of B which corresponds to g_a will be denoted A_a. It is monotonely decreasing with a, and if (1) holds then

$$\omega_{A_a} = 1 \quad \text{for} \quad a < p, \qquad \omega_{A_a} = 0 \quad \text{for} \quad q < a. \tag{23}$$

We now take any finite set of non-exceptional numbers $b_0 < b_1 < \cdots < b_n$, $b_0 < p$, $b_n > q$, and in relation

$$(b - a)\omega_{A_b} \leqq f(b) - f(a) \leqq (b - a)\omega_{A_a} \qquad (a < b)$$

we put $a = a_{\nu-1}$, $b = a_\nu$ and we sum over ν. Using relation (23) and the relation $f(a) = a$, $a < p$, and writing $E_\nu = A_{b_{\nu-1}} - A_{b_\nu}$ we hence obtain

$$\sum_{\nu=1}^{n} b_{\nu-1}\omega_{E_\nu} \leqq f \leqq \sum_{\nu=1}^{n} b_\nu \omega_{E_\nu} \tag{24}$$

and hence

$$\sum_{\nu=1}^{n} b_{\nu-1} \mid E_\nu \mid \leqq Mf \leqq \sum_{\nu=1}^{n} b_\nu \mid E_\nu \mid. \tag{25}$$

Since the maximal difference $b_\nu - b_{\nu-1}$ can be made arbitrarily small, relations (24) and (25) show that f belongs to the class of functions R_0 on $[B, vE]$.

The proof of the last part of theorem 3 runs as in "A," p. 775, bottom.

The reader will easily establish the following extension.

THEOREM 3′. *Theorem 3 also holds if the values of the functional* Mf (*and the measure* vE) *are elements of another partially ordered space* S *which has properties* 1) *and* 2), *in which every set which is bounded from above* (*below*) *has a* sup (inf) *and which satisfies the following condition*:

(i) Every concave function $\varphi(a)$, $p \leqq a \leqq q$, from real numbers to S has a derivative (in the sense of order) for a dense set of values a.

It is interesting to consider the following modification of this condition.

(ii) Every monotone function $x = f(a)$ is continuous for a dense set of values a.

(iii) The points a at which $f(a)$ is discontinuous have a potency smaller than that of the continuum.

(iv) The discontinuities of $f(a)$ are at most denumerable.

(v) The space S is complete in a Banach norm $\| x \|$ for which $x_1 > x_2 > \cdots > 0$, $x_0 = \inf (x_1, x_2, \cdots) > 0$ implies $\| x_1 \| > \| x_2 \| > \cdots \to \| x_0 \| > 0$.

Conditions (i) and (ii) are equivalent. This follows easily from the facts that for $\varphi(a)$ concave, the function

$$f(a) = \inf_{h>0} \frac{\varphi(a+h) - \varphi(a)}{h} \tag{26}$$

is monotonely decreasing, that for $f(a)$ decreasing the function

$$\varphi(a) = \varphi(0) + \int_0^a f(a)\, da \tag{27}$$

exists and is concave and that (26) implies (27). Conditions (iii) obviously implies (ii). However (ii) also implies (iii). In fact if (iii) does not hold then there exists a monotone function $g(a)$ such that

$$Dg(a) = \inf_{h>0} (g(a+h) - g(a-h))$$

is > 0 for values $\{a_\nu\}$, the index ν ranging over the continuum $0 < \nu < 1$. We now consider the functions

$$f_\nu(a) = \begin{cases} 0 & \text{for} \quad a < \nu \\ Dg(a_\nu) & \text{for} \quad a \geqq \nu \end{cases}$$

and at each point we define

$$f(a) = \sup_\nu f_\nu(a).$$

It is not hard to decide that $f(a)$ has the jump $Dg(a_\nu)$ at $a = \nu$ and thus violates (i). Condition (iv) implies condition (iii), and under the continuum hypothesis it is equivalent to it. Finally (v) implies (iv) and thus all other conditions. In fact, by our assumptions, $\| f(a) \|$ is increasing and $\| Df(a) \| = D \| f(a) \|$. Thus $Df(a) > 0$ for a countable number of points only.

PRINCETON UNIVERSITY.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

ON THE MEAN ERGODIC THEOREM

BY L. W. COHEN

(Received May 23, 1939)

The mean ergodic theorem of von Neumann[1] has been extended to spaces $L_p(p \geqq 1)$ by F. Riesz.[2] K. Yosida,[3] using a method similar to Riesz, obtains the following extension: *If T is linear on a Banach space B to B, the iterates T^n of T are bounded and*

$$L_n x = \frac{1}{n} \sum_{j=1}^{n} T^j x, \qquad n = 1, 2, \cdots,$$

is a weakly compact set, then there is an $x_0 \epsilon B$ such that $Tx_0 = x_0$ and $\lim L_n x = x_0$. The theorem may be regarded as stating[4] that the sequence $T^j x$ is transformed by a regular matrix[5] namely that of the Cesàro means of order 1, into a sequence $L_n x$ which, if it is weakly compact, converges strongly to an element invariant under T. A natural problem is that of determining a class of regular matrices a_{nj} such that if

$$L_n x = \sum_{j=1}^{\infty} a_{nj} T^j x$$

is weakly compact, then the sequence converges strongly to an $x_0 \epsilon B$ and $Tx_0 = x_0$. It will be shown that a sufficient condition for such a class is that

$$\lim_k \sum_{j=k}^{\infty} | a_{n,j+1} - a_{nj} | = 0$$

uniformly in n. In particular this is the case for the matrices of the Cesàro means of any positive order.

We first establish the

LEMMA. *If T and L_n, $n = 1, 2, \cdots$, are linear transformations on a Banach space B to B such that $TL_n = L_n T$ and, for some x,* $\lim L_n(x - Tx) = 0$ *and $L_n x$ converges weakly to x_0, then $Tx_0 = X_0$.*

---

[1] J. von Neumann, "Proof of the Quasi-Ergodic Hypothesis," *Proc. Nat. Acad. Sci.* vol. 18(1932) p. 70.

[2] F. Riesz, "Some Mean Ergodic Theorems," *Jour. Lond. Math. Soc.* vol. 13(1938) p. 274. Cf. E. Hopf, *Ergodentheorie*, Berlin, (1937) p. 23.

[3] K. Yosida, "Mean Ergodic Theorem in Banach Space," *Proc. Imp. Acad. Tokyo* vol. 14(1938) p. 292.

[4] The writer is indebted to Prof. J. D. Tamarkin for acquainting him with Yosida's note in his lectures and for discussing the general problem with him.

[5] A matrix a_{nj} is regular if for every convergent sequence of numbers ξ_j, $\lim_n \sum_{j=1}^{\infty} a_{nj}\xi_j = \lim_j \xi_j$.

PROOF. Let $\bar{x}$ be any linear functional on B. Then

$$\lim_n \bar{x}(L_n x - x_0) = 0 \tag{1}$$

from the definition of weak convergence. Since T is linear on B to B

$$\lim_n \bar{x}(TL_n x - Tx_0) = 0. \tag{2}$$

From $\lim L_n(x - Tx) = 0$ and the continuity of $\bar{x}$,

$$\lim_n \bar{x}(L_n x - L_n Tx) = 0. \tag{3}$$

Since T and L_n commute, we may write for all n

$$\bar{x}(x_0 - Tx_0) = \bar{x}(x_0 - L_n x) + \bar{x}(L_n x - L_n Tx) + \bar{x}(TL_n x - Tx_0).$$

Thus from 1, 2, 3 it follows that $\bar{x}(x_0 - Tx_0) = 0$ for all linear functionals on B. Hence $Tx_0 = x_0$.

THEOREM. *If T is linear on a Banach space B to B such that $\| T^n \| \leqq A (T^n = TT^{n-1})$, a_{nj} is a regular matrix such that*

$$\lim_k \sum_{j=k}^{\infty} | a_{n,j+1} - a_{nj} | = 0$$

uniformly in n and

$$L_n x = \sum_{j=1}^{\infty} a_{nj} T^j x$$

is a weakly compact set, then there is an $x_0 \epsilon B$ such that $Tx_0 = x_0$ and $\lim L_n x = x_0$.

PROOF. The Toeplitz conditions[6] for a regular matrix a_{nj} are

$$\sum_{j=1}^{\infty} | a_{nj} | \leqq M, (n = 1, 2, \cdots); \quad \lim_n a_{nj} = 0, (j = 1, 2, \cdots); \quad \lim_n \sum_{j=1}^{\infty} a_{nj} = 1. \tag{1}$$

Since B is complete and for $z \epsilon B$

$$\left\| \sum_{j=s}^{t} a_{nj} T^j z \right\| \leqq A \| z \| \sum_{j=s}^{t} | a_{nj} |$$

the L_n are defined on B. Further $\| L_n \| \leqq AM$ since

$$\| L_n z \| \leqq \sum_{i=1}^{\infty} | a_{nj} | \, \| T^j z \| \leqq A \| z \| \sum_{j=1}^{\infty} |a_{nj}| \leqq AM \| z \|. \tag{2}$$

Also

$$TL_n z = \sum_{j=1}^{\infty} a_{nj} T^{j+1} z = L_n Tz. \tag{3}$$

---

[6] O. Toeplitz, "Über allgemeine lineare Mittelbildungen," *Prace. Mat. Fiz.* vol. 22(1911) p. 113.

Now for any $z \in B$

$$\| L_n(z - Tz) \| = \left\| a_{n1} Tz + \sum_{j=1}^{\infty} (a_{n,j+1} - a_{nj}) T^{j+1} z \right\|$$

$$\leqq A \| z \| \left(| a_{n1} | + \sum_{j=1}^{k-1} | a_{n,j+1} - a_{nj} | + \sum_{j=k}^{\infty} | a_{n,j+1} - a_{nj} | \right)$$

for all n and $k > 1$. Given $\epsilon > 0$ there is, by the hypothesis of uniformity, a k_ϵ such that

$$\| L_n(z - Tz) \| \leqq A \| z \| \left(2 \sum_{j=1}^{k_\epsilon} | a_{nj} | + \epsilon \right), \qquad n = 1, 2, \cdots.$$

From 1 there is an n_ϵ such that

$$\sum_{j=1}^{k_\epsilon} | a_{nj} | < \epsilon, \qquad n > n_\epsilon .$$

Thus

$$\| L_n(z - Tz) \| \leqq 3A \| z \| \epsilon, \qquad n > n_\epsilon . \tag{4}$$

Since $L_n x$ is weakly compact there is an $x_0 \in B$ and a subsequence L_{n_i} such that $L_{n_i} x$ converges weakly to x_0. Thus because of 3 and 4 the T and L_{n_i} satisfy the conditions of the lemma and

$$Tx_0 = x_0 . \tag{5}$$

Let B_0 be the closed linear manifold determined by the set of $z - Tz$ for all $z \in B$. We show that $x_0 - x \in B_0$. If not, there is a linear functional $\bar{x}_0$ on B such that

$$\bar{x}_0 u = 0, \qquad u \in B_0 ; \qquad \bar{x}_0(x_0 - x) = 1. \tag{6}$$

Now $T^k x - T^{k+1} x \in B_0$, $k = 0, 1, 2, \cdots$, so that $\bar{x}_0(T^k x - T^{k+1} x) = 0$. Since $x - Tx \in B$ and

$$\bar{x}_0(x - T^j x) = \bar{x}_0(x - T^{j-1} x) + \bar{x}_0(T^{j-1} x - T^j x) = \bar{x}_0(x - T^{j-1} x)$$

we have

$$\bar{x}_0 x = \bar{x}_0 T^j x, \qquad j = 1, 2, \cdots . \tag{7}$$

From 4 and $\| L_n \| \leqq AM$ by 2, one sees that

$$\lim_n L_n u = 0, \qquad u \in B_0 . \tag{8}$$

Since $\bar{x}_0$ is linear we get from 1 and 7

$$\bar{x}_0 L_n x = \sum_{j=1}^{\infty} a_{nj} \bar{x}_0 T^j x = \left(\sum_{j=1}^{\infty} a_{nj} \right) \bar{x}_0 x, \qquad n = 1, 2, \cdots,$$

$$\lim_n \bar{x}_0 L_n x = \bar{x}_0 x. \tag{9}$$

Then from the weak convergence of $L_{n_i}x$ to x_0 and 9

$$0 = \lim_i \bar{x}_0(x_0 - L_{n_i}x) = \bar{x}_0 x_0 - \lim_i \bar{x}_0 L_{n_i} x = \bar{x}_0 x_0 - \bar{x}_0 x = \bar{x}_0(x_0 - x)$$

contrary to 6. Hence $x_0 - x \in B$.

Now from 5, $T^j x_0 = x_0$, $j = 1, 2, \cdots$ and so

$$L_n x_0 = \sum_{j=1}^{\infty} a_{nj} T^j x_0 = \left(\sum_{j=1}^{\infty} a_{nj}\right) x_0,$$

$$\lim_n L_n x_0 = x_0 \tag{10}$$

because of 1. Writing $x = x_0 + (x - x_0)$ we find from $x - x_0 \in B_0$, 8 and 10

$$\lim L_n x = x_0$$

and the theorem is proved.

The Cesàro matrix

$$a_{nj} = \frac{\binom{r+j-1}{j}}{\binom{r+n}{n}}, \qquad r > 0, \quad 0 \leqq j \leqq n, \quad n = 1, 2, \cdots,$$

satisfies the uniformity condition since $a_{nj} = 0$ for $j > n$ and

$$\begin{aligned} &| a_{n0} | + \sum_{j=0}^{n} | a_{n,j+1} - a_{nj} | \\ &\quad = \frac{1}{\binom{r+n}{n}} \left\{ 1 + \sum_{j=0}^{n-1} \left[\binom{r+j}{j+1} - \binom{r+j-1}{j} \right] + \binom{r+n-1}{n} \right\} \\ &\quad = 2 \frac{\binom{r+n-1}{n}}{\binom{r+n}{n}} < \epsilon, \qquad n > r_\epsilon . \end{aligned}$$

COROLLARY. *If T is linear on B to B, $\| T^n \| \leqq A$ and*

$$L_n x = \frac{1}{\binom{r+n-1}{n-1}} \sum_{j=0}^{n-1} \binom{r+j-1}{j} T^{n-j} x, \qquad r > 0,$$

is weakly compact, then there is an x_0 such that $Tx_0 = x_0$ and

$$\lim_{n-m\to\infty} L_{n,m} x = \lim_{n-m\to\infty} \frac{1}{\binom{r+n-m-1}{n-m-1}} \sum_{j=0}^{n-m-1} \binom{r+j-1}{j} T^{n-j} x = x_0.$$

PROOF. Since the conditions of the theorem are satisfied there is an $x_0 \epsilon B$ such that $Tx_0 = x_0$ and $\lim L_n x = x_0$. But $L_{n,m}x = T^m L_{n-m}x$; $T^m x_0 = x_0$; $L_{n-m}x_0 = x_0$. Thus

$$\| L_{n,m}x - x_0 \| = \| T^m L_{n-m}(x - x_0) \| \leqq A \| L_{n-m}(x - x_0) \| \leqq A\epsilon,$$
$$n - m > N_\epsilon .$$

BROWN UNIVERSITY.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

ON INTERPOLATION. III. INTERPOLATORY THEORY OF POLYNOMIALS

BY PAUL ERDÖS AND PAUL TURÁN[1]

(Received April 27, 1939)

Dedicated to Professor L. Fejér on the occasion of his sixtieth birthday

This paper may be read without a knowledge of our first two papers on interpolation.

Let

$$\mathfrak{M} \equiv \left\{ \begin{array}{llll} x_1^{(1)} & & & \\ x_1^{(2)}, & x_2^{(2)} & & \\ \vdots & & & \\ x_1^{(n)}, & x_2^{(n)}, & \cdots, & x_n^{(n)} \\ \vdots & \vdots & & \ddots \end{array} \right\} \tag{1}$$

be a triangular matrix where for each n

$$1 \geqq x_1^{(n)} > x_2^{(n)} > \cdots > x_n^{(n)} \geqq -1. \tag{2}$$

All the $x_\nu^{(n)}$ may be written in the form $\cos \vartheta_\nu^{(n)}$; hence for each $\mathfrak{M}$ we may define a triangular matrix

$$\mathfrak{M}' \equiv \left\{ \begin{array}{llll} \vartheta_1^{(1)} & \ddots & & \\ \vdots & & & \\ \vartheta_1^{(n)}, & \vartheta_2^{(n)}, & \cdots, & \vartheta_n^{(n)} \\ \vdots & \vdots & & \ddots \end{array} \right\} \tag{3}$$

with

$$0 \leqq \vartheta_1^{(n)} < \cdots < \vartheta_n^{(n)} \leqq \pi. \tag{4}$$

Let $f(x)$ be defined in $[-1, +1]$; then we define the n^{th} Lagrange interpolatory polynomial of $f(x)$ with respect to $\mathfrak{M}$ as the polynomial $L_n(f)$ of degree $(n-1)$ at most taking at the points $x_1^{(n)}, x_2^{(n)}, \cdots, x_n^{(n)}$ the values $f(x_1^{(n)}), f(x_2^{(n)}), \cdots, f(x_n^{(n)})$. It may be verified that

$$L_n(f) \equiv \sum_{\nu=1}^{n} f(x_\nu^{(n)}) l_{\nu,n}(x) \equiv \sum_{\nu=1}^{n} f(x_\nu) l_\nu(x) \equiv \sum_{\nu=1}^{n} f(x_\nu^{(n)}) \frac{\omega_n(x)}{\omega_n'(x_\nu^{(n)})(x - x_\nu^{(n)})}, \tag{5a}$$

[1] Some of these results have been presented before the Mathematical Association in Budapest in April, 1937.

where

$$\omega_n(x) \equiv \prod_{\nu=1}^{n} (x - x_\nu^{(n)}) \equiv \prod_{\nu=1}^{n} (x - x_\nu). \tag{5b}$$

We shall explicitly indicate the upper and double indices only when some misunderstanding may arise. The polynomials $l_\nu(x)$ (for which we omitted to indicate explicitly their dependence upon n) are independent of $f(x)$ and dependent only upon $\mathfrak{M}$, ν, and n; following Fejér they are called the fundamental functions of interpolation. For these it is easy to verify that

$$\sum_{\nu=1}^{n} l_{\nu,n}(x) \equiv \sum_{\nu=1}^{n} l_\nu(x) \equiv 1, \qquad n = 1, 2, \cdots, \tag{6a}$$

and more generally

$$L_n(r) \equiv \sum_{\nu=1}^{n} r(x_\nu) l_\nu(x) \equiv r(x), \qquad n = k+1, k+2, \cdots \tag{6b}$$

where $r(x)$ denotes any polynomial of degree k. The numbers $\int_{-1}^{+1} l_\nu(x)\,dx \equiv \lambda_\nu^{(n)} \equiv \lambda_\nu$ (depending only upon $\mathfrak{M}$, ν, and n) are called the Cotes numbers belonging to $\mathfrak{M}$. From (6a) we evidently have

$$\sum_{\nu=1}^{n} \lambda_\nu^{(n)} \equiv \sum_{\nu=1}^{n} \lambda_\nu = 2. \tag{7}$$

We intend to consider chiefly the case of two general and very often used matrices. The first of them is obtained as follows: let $p(x)$ be nonnegative and integrable in Lebesgue's sense (L-integrable) for $[-1, +1]$. Then a sequence of uniquely determined polynomials $\omega_0(x)$, $\omega_1(x)$, $\cdots$, corresponds to $p(x)$ so that $\omega_n(x)$ is a polynomial of degree n with

$$\text{coeff. } x^n \text{ in } \omega_n(x) = 1 \tag{8a}$$

and

$$\int_{-1}^{1} \omega_n(x)\omega_m(x)p(x)\,dx = 0, \qquad n \neq m. \tag{8b}$$

The sequence of such polynomials is called orthogonal with respect to the weight function $p(x)$. The sequence of polynomials $\Omega_0(x)$, $\cdots$, $\Omega_n(x)$ $\cdots$, for which

$$\int_{-1}^{1} \Omega_n(x)\Omega_m(x)p(x)\,dx = 0 \qquad n \neq m, \tag{8c}$$

$$\int_{-1}^{1} \Omega_n(x)^2 p(x)\,dx = 1 \qquad n = 0, 1, \cdots, \tag{8d}$$

$$\text{coeff. } x^n \text{ in } \Omega_n(x) \text{ is greater than } 0, \qquad n = 0, 1, \cdots, \tag{8e}$$

we call a sequence of normal-orthogonal polynomials with respect to $p(x)$. The polynomials $\omega_n(x)$ and $\Omega_n(x)$ evidently differ only in a constant factor dependent only upon n. By (8b) it is easy to see that all roots of $\omega_n(x)$ are real and situated in $[-1, +1]$. Taking these roots for $n = 1, 2, \cdots$ we obtain the so called p-matrix. It is well known that for $p(x) \equiv 1$ we obtain the sequence of Legendre-polynomials $P_n(x)$, for $p(x) = 1/\sqrt{(1-x^2)}$ the Tchebycheff-polynomials $T_n(x)$, and in general, for $p(x) = (1-x)^\alpha(1+x)^\beta$ with $\alpha > -1$, $\beta > -1$, the Jacobi-polynomials $P_n^{(\alpha,\beta)}(x)$.

The second class of matrices has been found by Fejér[2] in his paper about Lagrange-interpolation. According to his notation the matrix $\mathfrak{M}$ is *normal*, if

$$\text{(9a)} \qquad v_k(x) \equiv 1 - \frac{\omega_n''(x_k)}{\omega_n'(x_k)}(x - x_k) \geqq 0,$$

$$-1 \leqq x \leqq +1, \qquad k = 1, 2, \cdots n, \qquad n = 1, 2, \cdots,$$

and it is *strongly normal*, if

$$\text{(9b)} \qquad 1 - \frac{\omega_n''(x_k)}{\omega_n'(x_k)}(x - x_k) \geqq c_1,$$

$$-1 \leqq x \leqq +1, \qquad k = 1, 2, \cdots n, \qquad n = 1, 2, \cdots,$$

where c_1—and later all the other c's—are positive constants independent of x, n, k. Their dependence upon accidental parameters will always be explicitly stated. Fejér proved that e.g. the sequence of Jacobi-polynomials $P_n^{(\alpha,\beta)}(x)$ presents a normal matrix if $-1 < \alpha \leqq 0$, $-1 < \beta \leqq 0$ and a strongly normal one, if $-1 < \alpha < 0$, $-1 < \beta < 0$. For this second matrix class, by

$$\text{(10)} \qquad \sum_{k=1}^{n} v_k(x) l_k(x)^2 \equiv 1,$$

we have

$$\text{(11)} \qquad | l_k(x) | \leqq \frac{1}{\sqrt{c_1}}$$

$$-1 \leqq x \leqq +1, \qquad k = 1, 2, \cdots n, \qquad n = 1, 2, \cdots$$

Orthogonal polynomials, and especially Jacobi-polynomials, play a most important part in many problems of analysis; we mention here only the works of Legendre, Laplace, Jacobi, Bruns Tchebycheff, A. Markoff, Stieltjes, Christoffel, Darboux, Fejér, S. Bernstein and Szegö. In the general theory of orthogonal polynomials (i.e. for general $p(x)$) an important step has been made by G. Szegö.[3] He succeeded in proving for a general class of weight-functions the asymptotic formulae of Laplace-Darboux concerning Jacobi-polynomials. Thus he proved

[2] L. Fejér: *Lagrangesche interpolation und die zugehörigen konjugierten Punkte*, Math. Ann., 1932, pp. 1–55.

[3] G. Szegö: *Über die Entwicklung einer analytischen Function usw.*, Math. Ann., 1921, pp. 188–212.

that if $p(x)$ is such that to $p(\cos\vartheta)\mid\sin\vartheta\equiv p_1(\vartheta)$ there exists a function $D(z)$ regular in $|z|<1$, here $\neq 0$ and for almost all ϑ

$$\lim_{r\to 1-0}|D(re^{i\vartheta})|^2 = p_1(\vartheta), \tag{12}$$

then in $|z|\geqq R>1$ for $n\to\infty$ uniformly

$$\lim_{n\to\infty}\frac{\Omega_n\left[\left(z+\frac{1}{z}\right)\frac{1}{2}\right]}{z^n}=\frac{1}{(2\pi)^{\frac{1}{2}}D\left(\frac{1}{z}\right)},$$

which determines the asymptotic behavior of the polynomials for any point of the z plane not lying in $[-1, +1]$. (12) is satisfied, if $p(x)\geqq 0$, further if $p(x)$ and $\log p(x)$ are Lebesgue-integrable in $[-1, +1]$. Further—and this is a deeper result—Szegö[4] gave for the fundamental interval itself i.e. for $[-1, +1]$ an asymptotic formula

$$\Omega_n(\cos\vartheta_0)\sim\left(\frac{2}{\pi\sin\vartheta_0\,p(\cos\vartheta_0)}\right)^{\frac{1}{2}}\cos\left[\left(n+\frac{1}{2}\right)\vartheta_0-\frac{\pi}{4}-\alpha\right], \tag{13}$$

where α depends in a given way upon $p(x)$. In order to give a simple example, he proved this for $\epsilon\leqq\vartheta_0\leqq\pi-\epsilon$ and $n\to\infty$, if $p(x)$ remains in $[-1, +1]$ between two positive bounds and the first and second derivatives of $p(x)$ in the same interval exist. S. Bernstein[5] proved a theorem, which is analogous to the above mentioned theorem of Szegö. He proved the asymptotic formula (13) if, in $[-1, +1]$ $p(x)\sqrt{(1-x^2)}$ remains between two positive bounds and uniformly satisfies here a logarithmic Lipschitz condition with the exponent $1+\epsilon$. For this theorem, Szegö gives a very simple proof in his book to be published. The papers of J. Shohat[6] also contain general results of this kind.

The problems concerning orthogonal polynomials can be divided into four classes: a) the behavior of the polynomials within the interval $[-1, +1]$ ((*internal behavior*), b) the behavior of the polynomials upon the plane cut along $[-1, +1]$ (*external behavior*), c) distance of consecutive roots (problems of the *finer distribution* of roots), d) number of roots in a fixed subinterval (problems of the *mean-distribution* of roots). Problems concerning a) are completely solved by Szegö and Bernstein for a rather general class of weight-functions; if we require the weight function only to satisfy

$$p(x)\geqq c_2 \qquad -1\leqq x\leqq +1, \tag{14a}$$

$$p(x) \text{ is Lebesgue-integrable in } [-1, +1], \tag{14b}$$

---

[4] G. Szegö: *Über den asymptotischen Ausdruck von Polynomen, die durch eine Orthogonälitatseigenschaft definiert sind,* Math. Ann., 1922, pp. 114–139.

[5] S. Bernstein: *Sur les polynomes orthogonaux on a segment fini,* Journal de Mathematiques, pp. 127–177.

[6] See J. Shohat: *Théorie générale des polynomes orthogonaux de Tchebichef,* Mémorial des Sciences Mathématiques, Fasc. LXVIII.

then Shohat[6] gives an upper estimate for the orthogonal polynomials belonging to $p(x)$. As far as we know there are no other general results in this direction. Question b) is settled by Szegö for rather general weights. As for c) we obtain from Szegö's formula that if the weight function throughout a subinterval has derivatives of the first and second order and remains throughout $[-1, +1]$ between two positive bounds (Szegö gives some other, more general condition), then for the n^{th} fundamental points $x_\nu^{(n)} = \cos \vartheta_\nu^{(n)}$ lying in that subinterval, we have $\lim_{n\to\infty} n(\vartheta_{\nu+1}^{(n)} - \vartheta_\nu^{(n)}) = \pi$; we obtain the same result for consecutive fundamental points from Bernstein's theorem, if throughout the interval $[-1, +1]$

$$c_3 \geqq p(x)\sqrt{(1 - x^2)} \geqq c_4 \tag{15a}$$

and if for $p(\cos \vartheta) \sin \vartheta = t(\vartheta)$ throughout and uniformly in $[0, \pi]$

$$|t(\vartheta + h) - t(\vartheta)| < \frac{c_5}{\log^{1+\epsilon} \frac{1}{h}} \tag{15b}$$

Concerning d) Szegö[7] implicitly proved, that, if the weights are non-negative in $[-1, +1]$, and if in the same interval $p(x)$ and $\frac{1}{\sqrt{(1 - x^2)}} \log p(x)$ are L-integrable, then the distribution of the n^{th} fundamental points is uniform, which means, that if $0 \leqq \alpha < \beta \leqq \pi$, then with $x_\nu^{(n)} = \cos \vartheta_\nu^{(n)}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_\nu 1 = \frac{\beta - \alpha}{\pi}, \qquad \alpha \leqq \vartheta_\nu^{(n)} \leqq \beta. \tag{16}$$

Szegö's and Bernstein's methods are based upon asymptotic formulae for polynomials. But it is probable that in the general case such a formula does not exist not even for continuous weights remaining between two positive bounds. Thus, in this way we cannot obtain any answer to questions such as e.g. what is the effect of the singularities of the weight function (loci of discontinuity, infinities, zeros) upon the distribution of roots, whether this effect is only local etc. The investigation of this last question will be a main object of our paper. Here we make use of a principle introduced by Fejér: we derive the structure of the matrix from the properties of interpolatory fundamental functions belonging to $\mathfrak{M}$. Fejér deals with two such properties. The first[2] is the property of being strongly normal, from which he deduces the relation

$$\lim_{n\to\infty} \max_{\nu=1,2,\cdots(n-1)} (x_{\nu+1}^{(n)} - x_\nu^{(n)}) = 0, \tag{17}$$

which—from what precedes—means a statement about the distribution of roots of certain Jacobi-polynomials. The second property[7] is the non-negativeness

[7] L. Fejér: *Mechanische Quadraturen mit positiven Colesschen Zahlen*, Math. Zeitschr., 1933, pp. 287-310. His proof gives also the following result: if there exists for the matrix

of the Cotes-numbers belonging to $\mathfrak{M}$, from which we once more obtain (17). Theorems deducing properties of $\mathfrak{M}$ from some interpolatory properties we shall call Fejérian theorems. We proved[8] two such theorems, the application of which to p-matrices gave the following two theorems.

I. *If throughout* $[-1, +1]$ $c_6 \leqq p(x) \leqq c_7$ *and* $p(x)$ *is L-integrable, then*

$$\vartheta_{\nu+1}^{(n)} - \vartheta_\nu^{(n)} \leqq \frac{c_8}{n}$$

$$\nu = 0, 1, 2, \cdots (n-1), n, \qquad n = 1, 2, \cdots, \qquad \vartheta_0^{(n)} = 0, \qquad \vartheta_{n+1}^{(n)} = \pi$$

for the corresponding matrix $\mathfrak{M}'$. (See (3).)

II. *If throughout the interval* $[-1, +1]$ $p(x) \geqq 0$, $p(x)$ *and* $\dfrac{1}{p(x)}$ *L-integrable, then*

$$\vartheta_{\nu+1}^{(n)} - \vartheta_\nu^{(n)} \leqq \frac{c_9 \log (n+1)}{n}$$

$$\nu = 0, 1, \cdots n. \qquad n = 1, 2, \cdots.$$

By systematic application of Fejér's principle we obtain Fejérian theorems for each of the four classes mentioned above, theorems, which may be applied to p-matrices as well as to strongly normal ones. Properly speaking we deduce the theory of both classes of polynomials from that of a more general class of polynomials, the roots of which form a matrix $\mathfrak{M}$, and for which the values of the fundamental functions $l_\nu(x)$ satisfy certain conditions.

In §2 we consider problem a). That will be the only section in which we shall not explicitly express a Fejérian-theorem. Our theorem I asserts for strongly normal polynomials

$$|\omega_n(x)| \leqq \frac{8}{\sqrt{c_1}} \frac{\sqrt{n}}{2^n}, \qquad -1 \leqq x \leqq +1, \qquad n = 1, 2, \cdots$$

where c_1 is any constant for which (11) is valid. This result cannot be essentially improved in $[-1, +1]$, but it is probable, that in $[-1+\epsilon, 1-\epsilon]$ the factor with $\sqrt{n}$ can be omitted and the factor $8/\sqrt{c_1}$ replaced by a $c_{10} = c_{10}(c_1, \epsilon)$. Theorem II applies to the orthogonal polynomials and it states that,

---

$\mathfrak{M}$ a function $s(x)$, non-negative and L-integrable on $[-1, +1]$, positive in $[a, b]$ and such that

$$\int_{-1}^{1} l_{\nu n}(x) s(x)\, dx \geqq 0, \qquad \nu = 1, 2, \cdots, n, \quad n = 1, 2, \cdots$$

then (17) holds in $[a, b]$. This is satisfied e.g. if the matrix $\mathfrak{M}$ is a p-matrix and $s(x) \equiv p(x)$; hence the roots of the polynomials orthogonal to a weight-function, non-negative and L-integrable on $[-1, +1]$ and positive throughout the subinterval $[a, b]$ cover the interval $[a, b]$ everywhere densely.

[8] P. Erdös and P. Turán: *On Interpolation II*, Annals of Math. 1938, pp. 703–724.

if the weight-function is non-negative and L-integrable in $[-1, +1]$ and if throughout the subinterval $[a, b]$ $p(x) \geqq m > 0$, then

$$(18a) \qquad |\omega_n(x)| < \left[\frac{72}{(b-a)m}\int_{-1}^{1} p(t)\,dt\right]^{\frac{1}{2}} \frac{n}{2^n} \quad \text{in} \quad a \leqq x \leqq b,\ n = 1, 2, \dots,$$

$$(18b) \qquad |\omega_n(x)| < \left[\frac{12}{m[\epsilon(b-a)]^{\frac{1}{2}}}\int_{-1}^{1} p(t)\,dt\right]^{\frac{1}{2}} \cdot \frac{\sqrt{n}}{2^n}$$

$$\text{in} \quad a + \epsilon \leqq x \leqq b - \epsilon,\ n = 1, 2, \dots.$$

For the case $a = -1, b = +1$ these estimations have been presented by Shohat. By the same method we obtain lower estimates for the orthogonal polynomials $\omega_n(x)$. More exactly: if the weight $p(x)$ is non-negative and L-integrable throughout $[-1, +1]$ and $p(x) \geqq m > 0$ throughout a subinterval $[a, b]$, then for any x (real or complex)

$$(19a) \qquad |\omega_n(x)| \geqq \left[c_{11}\frac{m}{(b-a)\displaystyle\int_{-1}^{1} p(t)\,dt}\right]^{\frac{1}{2}} \left(\frac{b-a}{4}\right)^n |x - x_\nu^{(n)}|,$$

where $x_\nu^{(n)}$ denotes the n^{th} fundamental point nearest to x. As a matter of fact this has importance only for the interval $[-1, +1]$. If in addition to the above properties throughout a subinterval $[c, d]$ of $[a, b]$ the weight is bounded, $p(x) \leqq M$, then a factor with $\sqrt{n}$ can be appended to the right side, if we take $c_{11} = c_{11}(M)$. For $c = a = -1, d = b = +1$ we find implicitly and qualitatively the same as Shohat.[9] By this and by the results of §3 we obtain e.g. that if for the L-integrable weight function $p(x) \geqq m > 0$ throughout $[-1, +1]$ and if $p(x) \leqq M$ throughout the subinterval $[e, f]$, then $\omega_n(x)$ takes in any $[x_{\nu+1}^{(n)}, x_\nu^{(n)}]$ lying in $[e + \epsilon, f - \epsilon]$ a value, greater than

$$(19b) \qquad \frac{c_{12}(\epsilon, M, m, e, f)}{2^n\sqrt{n}}$$

It is probable, that in (18a) and (18b) the factor n or $\sqrt{n}$ may be improved to $c_{13}(a, b, m)\sqrt{n}$ or to $c_{14}(\epsilon, a, b, m)$ respectively—this is true in the mean—and also in (19b) we may omit from the denominator the factor with $\sqrt{n}$. If in $[-1, +1]$ $p(x) \geqq m > 0$ and in the subinterval $[a, b]$ $p(x) \leqq M$, then we proved that there exists an $\eta(n)$ such that $\eta(n) \to 0$ for $n \to \infty$ and that in $[a + \epsilon, b - \epsilon]$

$$|\omega_n(x)| < c_{15}(a, b, \epsilon, m, M)\frac{\eta(n)\sqrt{n}}{2^n}.$$

We omit the details of the proof.

---

[9] See footnote 6, p. 41, formula (60).

In §3 we are concerned with b) problems. The base of the investigation is the following Fejérian theorem: If for the matrix $\mathfrak{M}$ for every $\epsilon > 0$

$$(20)\qquad [|\, l_k(x)\, |]^{1/n} \leqq 1 + \epsilon, \qquad n > c_{16}(\epsilon), \quad k = 1, 2, \cdots n, \quad -1 \leqq x \leqq +1,$$

then for any fixed z of the complex plane cut along $[-1, +1]$ we have

$$(21)\qquad \lim_{n\to\infty} [\omega_n(z)]^{1/n} = \frac{z + \sqrt{(z^2 - 1)}}{2},$$

where we are to take those values of the n^{th} and square roots, which are positive on the positive real axis. Condition (20) is abundantly satisfied for sequences of strongly normal polynomials. Thus the asymptotic representation (21) applies for these too.

We shall see that if in $[-1, +1]$ the L-integrable $p(x)$ is $\geqq 0$ and its roots form an aggregate of measure 0, then (20) is satisfied hence (21) holds too. Formula (21) presents less than the above quoted formula of Szegö but it refers to a wider class of weight-functions: e.g. (21) holds for the weight-function $p(x) = e^{-1/x^2}$, whereas Szegö's formula has nothing to say in this case.

We shall give a direct and elementary proof of the aforesaid theorem, but we are bound to mention that it is to be deduced indirectly from a deep theorem of L. Kalmár[10] by the following note of Polya:[11] If upon the matrix $\mathfrak{M}$ we have uniformly in $[-1, +1]$

$$\lim_{n\to\infty} [|\, l_1(x)\, | + \cdots + |\, l_n(x)\, |]^{1/n} = 1,$$

then the Lagrange parabolas taken upon $\mathfrak{M}$ of a function $f(x)$ analytic in this interval uniformly converge to $f(x)$. In order to prove this note standard theorems about approximation of analytic functions are required.

On the other hand by a further theorem of Kalmár[10] it follows, that the ele-

---

[10] L. Kalmár: *Az interpolátióról*, Mathematikai és Fizikai Lapok, 1927, pp. 120-149 (Hungarian). This gives the following result: Let $\mathfrak{B}$ be the closed interior of Jordan-curve l on the complex z-plane and let $x = \varphi(z)$ be regular on the exterior of l and continuous on the closed exterior of l, which maps $\mathfrak{B}$ upon the exterior of a circle $|\, x\, | \leqq c$ with $\lim_{|z|\to\infty} \frac{\varphi(z)}{z} = 1$. Let the matrix $\mathfrak{M}$ be given in $\mathfrak{B}$ and $\omega_n(z) = \prod_{\nu=1}^{n} (z - z_\nu^{(n)})$. Then a necessary and sufficient condition that $\lim_{n\to\infty} L_n(f) = f(z)$ uniformly in $\mathfrak{B}$ for any $f(z)$ regular in $\mathfrak{B}$ is that $\lim_{n\to\infty} [\omega_n(z)]^{1/n} = \varphi(z)$ for any z of the exterior of l. We use this only in the case if l is the interval $[-1, +1]$.

[11] G. Polya: *Über die Konvergenz von Quadraturverfahren*, Math. Zeitschrift, 1933, pp. 264-287.

ments of $\mathfrak{M}'$ belonging to $\mathfrak{M}$ are uniformly distributed in $[0, \pi]$.[12] From this we incidentally obtained the following result: if $p(x)$ is non-negative and L-integrable in $[-1, +1]$, and further, if the roots of this $p(x)$ form an aggregate of measure 0, then the elements of the matrix $\mathfrak{M}'$ belonging to the roots of the respective orthogonal polynomials are uniformly distributed. (We can prove this result in a direct and elementary way, too.)

In §4 we consider c) problems. The basis of the general consideration is given by the following Fejérian-theorem: If a matrix $\mathfrak{M}'$ is such, that for a subinterval $[\alpha, \beta]$ of $[0, \pi]$ with

$$\vartheta_{\nu-1}^{(n)} < \alpha \leqq \vartheta_\nu^{(n)} < \vartheta_{\nu+1}^{(n)} < \cdots < \vartheta_\mu^{(n)} \leqq \beta < \vartheta_{\mu+1}^{(n)}$$

we have

$$| l_k(x) | \leqq K, \qquad k = \nu, \nu + 1, \cdots \mu, \qquad a = \cos \beta \leqq x \leqq \cos \alpha = b,$$

and, in the same subinterval, the absolute value of the other n^{th} fundamental functions does not exceed $c_{16}n^{c_{17}}$, then

$$\frac{[\epsilon(b-a)]^{\frac{1}{2}}}{K} \cdot \frac{1}{n} \leqq \vartheta_{k+1}^{(n)} - \vartheta_k^{(n)} \leqq \frac{c_{18}(c_{16}, c_{17}, \epsilon, a, b)K}{n}, \tag{22}$$

if $\vartheta_k^{(n)}$ and $\vartheta_{k+1}^{(n)}$ are in $[\alpha + \epsilon, \beta - \epsilon]$. If $[\alpha, \beta] \equiv [0, \pi]$, then c_{18} is independent of ϵ and the estimate holds for all $[\vartheta_k^{(n)}, \vartheta_{k+1}^{(n)}]$ $(k = 1, 2, \cdots, n-1)$, the upper estimation holds even, as we proved[8] for $k = 0$ and $k = n$, if $\vartheta_0^{(n)} = 0, \vartheta_{n+1}^{(n)} = \pi$. The content of the theorem may briefly be expressed as follows: if the fundamental functions belonging to the fundamental points of a subinterval are bounded and the other fundamental functions are in the same subinterval not excessively great, then the distribution of the matrix is approximately uniform in that subinterval. In our paper cited under[8] we already proved, that the estimate of the form (22) holds in the case of strongly normal polynomials for any pair $[\vartheta_k^{(n)}, \vartheta_{k+1}^{(n)}]$ with an absolute constant c_{18}. For orthogonal polynomials we obtain that, if the L-integrable weight function is non-negative in $[-1, +1]$ and if $0 < m \leqq p(x) \leqq M$ in a subinterval $[\cos \beta, \cos \alpha]$, then for any pair $[\vartheta_k^{(n)}, \vartheta_{k+1}^{(n)}]$ in $[\alpha + \epsilon, \beta - \epsilon]$ we have

$$\frac{c_{19}(m, M, \alpha, \beta, \epsilon)}{n} \leqq \vartheta_{k+1}^{(n)} - \vartheta_k^{(n)} \leqq \frac{c_{20}(m, M, \alpha, \beta, \epsilon)}{n}.$$

For $[\alpha, \beta] \equiv [0, \pi]$ c_{19} and c_{20} are independent of ϵ; in our paper we proved the upper estimate for this case, we omit the details of the lower estimate.

If in the subinterval $[\cos \beta, \cos \alpha]$ $p(x)\sqrt{(1-x^2)} \geqq m > 0$ and besides it

---

[12] This means of course, that for any fixed subinterval $[\alpha, \beta]$ of $[0, \pi]$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\substack{\nu \\ \alpha \leqq \vartheta_\nu^{(n)} \leqq \beta}} 1 = \frac{\beta - \alpha}{\pi} \text{ holds.}$$

$p(x)\sqrt{(1-x^2)}$ is continuous, very much more is to be said from the n^{th} fundamental-points situated in $[\alpha+\epsilon, \beta-\epsilon]$. In this case we have for $n \to \infty$

$$\vartheta_{k+1}^{(n)} - \vartheta_k^{(n)} \sim \frac{\pi}{n}.$$

The proof is also based upon the analysis of interpolatory forms but not upon a Fejérian-theorem; it is important to notice that the formula obtained for the fundamental functions relative to $[\alpha+\epsilon, \beta-\epsilon]$ is to some extent an asymptotic one. The interval $[\alpha+\epsilon, \beta-\epsilon]$ may be replaced by $\left[\alpha+\frac{A(n)}{n}, \beta-\frac{A(n)}{n}\right]$, where $A(n)$, though arbitrarily slowly, tends to infinity, and we may postulate other, more general conditions for the weight function. In the case of strongly normal polynomials the former of us proved in another way, that for $\frac{A(n)}{n} \leqq \vartheta_k^{(n)} < \vartheta_{k+1}^{(n)} \leqq \pi - \frac{A(n)}{n}$ the difference $\vartheta_{k+1}^{(n)} - \vartheta_k^{(n)} \sim \frac{\pi}{n}$. We do not give the details of the proof.

In §5 we consider d) problems. The analysis is based upon two Fejérian-theorems. The first of them states, that the uniform distribution in the sense (16) of the matrix $\mathfrak{M}$ is a consequence of condition (20); we give for this a completely elementary direct proof. If for a matrix $\mathfrak{M}$ with the absolute constant K'

$$|\, l_\nu(x)\,| \leqq K', \qquad -1 \leqq x \leqq +1, \qquad \nu = 1, 2, \cdots n$$

then more exactly

$$(23) \quad -c_{22}(K',\epsilon)\{(\beta-\alpha)n\}^{\frac{1}{2}+\epsilon} < \sum_{\substack{\nu \\ \alpha \leqq \vartheta_\nu^{(n)} \leqq \beta}} 1 - \frac{\beta-\alpha}{\pi} n < \{(\beta-\alpha)n\}^{\frac{1}{2}+\epsilon} c_{22}(K',\epsilon)$$

for $(\beta-\alpha)n > c_{23}(K', \epsilon)$. This means, that for uniformly bounded fundamental functions the uniform distribution is already effected for very small subintervals $[\alpha, \beta]$, the size of which depend upon n. If $[\alpha, \beta]$ means any interval in $[0, \pi]$, then by the condition

$$(24) \quad |\, l_\nu(x)\,| \leqq c_{24} n^{c_{25}}, \qquad -1 \leqq x \leqq +1, \qquad \nu = 1, 2, \cdots n, \qquad n = 1, 2, \cdots$$

we have

$$(25) \qquad \left|\sum_{\substack{\nu \\ \alpha \leqq \vartheta_\nu^{(n)} \leqq \beta}} 1 - \frac{\beta-\alpha}{\pi} n\right| < c_{26}(c_{25}, c_{24}, \epsilon) n^{\frac{1}{2}+\epsilon},$$

which establishes the uniform distribution already for intervals of the length $1/n^{\frac{1}{2}-2\epsilon}$. This is not very much weaker than the former conclusion.

By applying the above-arguments to sequences of strongly normal polynomials we immediately see that the fundamental points are distributed according to (23). Thus for orthogonal polynomials we obtained a new and strictly

elementary proof of our theorem that, if the L-integrable $p(x)$ weight-function is in $[-1, +1]$ not less than 0 and the aggregate of the points x with $p(x) = 0$, is of measure 0, then the distribution of the elements of the matrix $\mathfrak{M}'$ formed of the roots of the respective orthogonal polynomials is uniformly dense in $[0, \pi]$. Here we must remark, that although our hypothesis is more general than that of Szegö, we obtained only the sufficient condition for the uniform distribution of the roots; the necessary and sufficient condition—as the first of us proved—is connected with the transfinite diameter of the aggregate of points, for which $p(x) = 0$. We omit the proof here.

Our second theorem states that, if the L-integrable weight function is $\geqq m > 0$ in $[-1, +1]$, then (25) holds for the corresponding matrix $\mathfrak{M}'$; if in addition, for $[-1, +1]$ $M \geqq p(x)\sqrt{(1 - x^2)} \geqq m$, then (23) holds too. If the aforesaid conditions are valid only for a subinterval and for the complementary subinterval of $[-1, +1]$ we postulate only the non-negativeness and the L-integrability, then nothing may be said with respect to the d) problems.

From the point of view of the theory of uniform distribution we make following remarks. Weyl's criterion for the uniform distribution of $\mathfrak{M}$ under (1) postulates, that for $n \to \infty$ the expressions $s_k \equiv \sum_{\nu=1}^{n} e^{2\pi i k \vartheta_\nu^{(n)}}$ tend to 0 for any positive integer k. Our theorems of §5 deduce the uniform distribution from the behavior of certain polynomials associated with $\mathfrak{M}$. It is to be noticed that instead of asymptotic equalities we have in the condition only inequalities and that we obtain also an error-term, that could not be obtained by Weyl's criterion. It would be plausible to ask, whether the uniform distribution with error-term is to be deduced from an inequality relative—in $[-1, +1]$—to $\omega_n(x)$ itself. The answer is affirmative; if in $[-1, +1]$ $|\omega_n(x)| \leqq \dfrac{A(n)}{2^n}$ with $A(n) \geqq 2$, then for a fixed subinterval $[\alpha, \beta]$ we have

$$\left| \sum_{\substack{\nu \\ \alpha \leqq \vartheta_\nu^{(n)} \leqq \beta}} 1 - \frac{\beta - \alpha}{\pi} n \right| \leqq 8[n \log A(n)]^{\frac{1}{2}}.$$

We will return to this problem on another occasion. If we disregard the error-term then, as we learned later the theorem is contained in a general theorem of Fekete[13] stating that the distribution of a matrix $\mathfrak{M}^+$ given upon any Jordan-curve l is uniform, if upon l the inequality $[|\omega_n(x)|]^{1/n} \leqq M$ holds, where M denotes the transfinite diameter of the Jordan-curve. Our argument essentially differs from his method.

From what is said before the reader may see the chief results of this paper: the uniformity of the method, the statement that the polynomials and their roots essentially depend only upon the local values of the weight function and asymptotic formulae of more general validity than before. We hope to consider the other fundamental problems in another paper.

---

[13] Oral communication.

1.

THEOREM I. *For strongly normal matrices we have in* $[-1, +1]$

$$|\omega_n(x)| \leqq \frac{8}{\sqrt{c_1}} \cdot \frac{\sqrt{n}}{2^n}, \qquad n = 1, 2, \cdots.$$

PROOF. As the arithmetic mean is not less than the geometric mean we may write

$$\frac{1}{n}\frac{1}{c_1} \geqq \frac{1}{n}\sum_{\nu=1}^{n} l_\nu(x)^2 \geqq \left[\prod_{\nu=1}^{n} l_\nu(x)^2\right]^{1/n} = \frac{|\omega_n(x)|^{2-2/n}}{\left[\prod_{\nu=1}^{n} \omega_n'(x_\nu)^2\right]^{1/n}}.$$

As the x_ν's are in $[-1, +1]$, we have after Schur[14]

$$\prod_{\nu=1}^{n} |\omega_n'(x_\nu)| < \frac{c_{28} n^{n+\frac{1}{2}}}{2^{n^2-2n}},$$

i.e.

$$|\omega_n(x)|^{2-2/n} < \frac{1}{c_1 n} \cdot \frac{c_{29} n^2}{2^{2n}},$$

hence

$$|\omega_n(x)| < \frac{c_{30}\sqrt{n}}{2^n}, \qquad \text{Q.e.d.}$$

This proof is very simple, but Schur's theorem which we applied is not of interpolatory nature. Hence it will perhaps be of some interest to give another proof for it. We require

LEMMA I. *If* $1 \geqq x_1^{(n)} > x_2^{(n)} > \cdots > x_n^{(n)} \geqq -1$, *then*

$$\sum_{\nu=1}^{n} \frac{1}{|\omega_n'(x_\nu)|} \geqq 2^{n-2}.$$

(Equality only for $\omega_n(x) = (x^2 - 1)U_{n-2}(x)$, where $U_k(\cos\vartheta) = \dfrac{1}{2^k}\dfrac{\sin(k+1)\vartheta}{\sin\vartheta}$, but for the present we shall not use this.)

PROOF. Let us fix in $[-1, +1]$ the values $\xi_1 > \xi_2 > \cdots > \xi_n$ and let us determine the polynomial $f(x)$ of degree $(n - 1)$, for which coeff. $x^{n-1} = 1$ and $\max_{\nu=1,2,\cdots n} |f(\xi_\nu)|$ is minimum. According to standard theorems such $f(x)$ exists and takes at the places ξ_ν with alternating signs the same absolute values $(\nu = 1, 2, \cdots, n)$. Thus by coeff. $x^{n-1} = 1$ we have

$$f(x) = \frac{\displaystyle\sum_{\nu=1}^{n} \frac{(-1)^{\nu+1}}{\omega'(\xi_\nu)} \cdot \frac{\omega(x)}{x - \xi_\nu}}{\displaystyle\sum_{\nu=1}^{n} \frac{(-1)^{\nu+1}}{\omega'(\xi_\nu)}},$$

[14] I. Schur: *Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten*, Math. Zeitschrift, 1918, pp. 377–402.

where $\omega(x) = \prod_{\nu=1}^{n}(x - \xi_\nu)$. The minimum value is given by the formula

$$M_n = \frac{1}{\sum_{\nu=1}^{n} \frac{(-1)^{\nu+1}}{\omega'(\xi_\nu)}} = \frac{1}{\sum_{\nu=1}^{n} \frac{1}{|\omega'(\xi_\nu)|}},$$

i.e.

$$\frac{1}{\sum_{\nu=1}^{n} \frac{1}{|\omega'(\xi_\nu)|}} = \min_{f=x^{n-1}+\cdots} \max_{\nu=1,2,\cdots n} |f(\xi_\nu)| \leqq \min_{f=x^{n-1}+\cdots} \max_{-1\leqq x\leqq+1} |f(x)| \leqq \frac{1}{2^{n-2}},$$

since for $f(x) = T_{n-1}(x)(T_{n-1}(\cos\vartheta) = \frac{1}{2^{n-2}}\cos(n-1)\vartheta)$, in $[-1, +1]$ $\max|T_{n-1}(x)| = \frac{1}{2^{n-2}}$. By taking the reciprocals we obtain the Lemma.

By the Lemma we immediately obtain that

$$|\omega_n(x)|\, 2^{n-2} \leqq |\omega_n(x)| \sum_{\nu=1}^{n} \frac{1}{|\omega_n'(\xi_\nu)|} \leqq 2 \sum_{\nu=1}^{n} \left| \frac{\omega_n(x)}{\omega_n'(\xi_\nu)(x - \xi_\nu)} \right|$$

$$= 2 \sum_{\nu=1}^{n} |l_\nu(x)| \leqq 2\sqrt{n} \left[\sum_{\nu=1}^{n} l_\nu(x)^2 \right]^{\frac{1}{2}} \leqq \frac{2}{\sqrt{c_1}} \sqrt{n},$$

which establishes the theorem. Notice that in both proofs we used only the fact that $\sum_{\nu=1}^{n} l_\nu(x)^2 < c_1$.

This result is not to be improved essentially in $[-1, +1]$, that is to be seen by the matrix given by the roots of the Jacobi-polynomial $P_n^{(-\epsilon,-\epsilon)}(x)$ for ϵ being any small fixed positive numbers. Its being strongly normal we already mentioned in the introduction. On the other hand by

$$P_n^{(\alpha,\beta)}(1) = \frac{2^n \binom{n+\alpha}{n}}{\binom{2n+\alpha+\beta}{n}}, \qquad \alpha > -1, \beta > -1$$

we have

$$P_n^{(\epsilon,-\epsilon)}(1) \sim c_{31}(\epsilon) \frac{n^{\frac{1}{2}-\epsilon}}{2^n}.$$

Theorem II. *If the L-integrable weight-function $p(x)$ is non-negative in $[-1, +1]$, and for the subinterval $[a, b] \geqq m(> 0)$, then in $[a, b]$*

$$|\omega_n(x)| \leqq \left[\frac{8}{(b-a)m} \int_{-1}^{1} p(t)\, dt \right]^{\frac{1}{2}} \cdot \frac{2n+1}{2^n},$$

whereas in $[a + \epsilon, b - \epsilon]$

$$|\omega_n(x)| \leqq 2 \left[\frac{1}{m[\epsilon(b-a-\epsilon)]^{\frac{1}{2}}} \int_{-1}^{1} p(t)\, dt \right]^{\frac{1}{2}} \cdot \frac{\sqrt{(2n+1)}}{2^n}.$$

Proof. As is known—and it may easily be verified—$\omega_n(x)$ minimizes the integral $J(f) \equiv \int_{-1}^{1} f(t)^2 p(t)\,dt$, if $f(t)$ runs over the polynomials of degree n with coeff. $x^n = 1$. Thus for $a \leqq x \leqq b$ we have

$$m \int_a^b \omega_n(t)^2\,dt \leqq \int_a^b \omega_n(t)^2 p(t)\,dt \leqq \int_{-1}^{1} \omega_n(t)^2 p(t)\,dt \leqq \int_{-1}^{1} T_n(t)^2 p(t)\,dt,$$

where $T_n(\cos\vartheta) = \frac{1}{2^{n-1}}\cos n\vartheta$. Hence

$$m \int_a^b \omega_n(t)^2\,dt \leqq \frac{4}{2^{2n}} \int_{-1}^{1} p(t)\,dt. \tag{26}$$

But then, according to a theorem of A. Markoff (stating that if for $a \leqq x \leqq b$ $|F(x)| \leqq M$, then here $|F'(x)| \leqq \frac{2M}{b-a} n^2$, where n denotes the degree of $F(x)$) for $a \leqq x \leqq b$ we have

$$|\omega_n(x)|^2 \leqq \frac{8}{2^{2n} m} \int_{-1}^{1} p(t)\,dt\, \frac{(2n+1)^2}{b-a}. \qquad \text{Q.e.d.}$$

By applying to (26) the theorem of Bernstein-Fejér (stating that if for $a \leqq x \leqq b$ $|F(x)| \leqq M$, then $|F'(x)| \leqq \frac{Mn}{[(b-x)(x-a)]^{\frac{1}{2}}}$, where n denotes the degree of $F(x)$) we obtain for $a + \epsilon \leqq x \leqq b - \epsilon$

$$\omega_n(x)^2 \leqq \frac{4}{m} \int_{-1}^{1} p(t)\,dt\, \frac{1}{[\epsilon(b-a-\epsilon)]^{\frac{1}{2}}} \cdot \frac{2n+1}{2^{2n}}. \qquad \text{Q.e.d.}$$

In connection with theorem II we mentioned that it is probable that the factor $\sqrt{n}$ in (18b) is to be improved to $c_{14}(\epsilon, a, b, m)$. This conjecture may to some extent be supported by the fact that from (26)

$$\left[\int_a^b \omega(t)^2\,dt\right]^{\frac{1}{2}} < \frac{2}{2^n}\left[\frac{1}{m}\int_{-1}^{1} p(t)\,dt\right]^{\frac{1}{2}}, \tag{27}$$

i.e. for $[a, b]$ the mean value of $|\omega_n(t)|$ is $O\left(\frac{1}{2^n}\right)$.

The proof of theorem II is very simple, but it is not of interpolatory character; thus we give a proof of such kind which with a slight modification gives the lower estimate indicated in the introduction, and besides it contains many elements needed in the following investigations.

Let the numbers

$$k_\nu^{(n)} = \int_{-1}^{1} l_{\nu,n}(t) p(t)\,dt \equiv k_\nu, \qquad \nu = 1, 2, \cdots, n, \qquad n = 1, 2, \cdots, \tag{28}$$

denote the Christoffel-numbers belonging to $p(x)$, then we have

LEMMA II. *In* $[-1, +1]$ *suppose* $p_1(x) \geqq p_2(x) \geqq 0$, *both L-integrable. If* $l_\nu(x)$ $(\nu = 1, 2, \cdots, n; n = 1, 2, \cdots)$ *stand for the fundamental functions and* k_ν *for the Christoffel-numbers belonging to* $p_1(x)$, $l_\nu^+(x)$ *and* k_ν^+ *for those belonging to* $p_2(x)$ *respectively, then for any fixed (real or complex)* x_0

$$\sum_{\nu=1}^{n} \frac{|l_\nu(x_0)|^2}{k_\nu} \leqq \sum_{\mu=1}^{n} \frac{|l_\nu^+(x_0)|^2}{k_\nu^+}, \qquad n = 1, 2, \cdots.$$

PROOF. Let x_0 denote any fixed number and determine the polynomial $F(x)$ of degree $(n - 1)$ at the utmost, for which $F(x_0) = 1$ and $I(F) \equiv \int_{-1}^{1} |F(t)|^2 p_1(t)\, dt$ is minimum.

We express $F(x)$ by the interpolatory polynomials belonging to the roots of n^{th} polynomial orthogonal to $p_1(x)$, then we have

$$F(x) = \sum_{\nu=1}^{n} d_\nu l_\nu(x),$$

i.e.

$$I(F) = \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} d_\mu \bar{d}_\nu \int_{-1}^{1} l_\mu(x) l_\nu(x) p_1(x)\, dx = \sum_{\nu=1}^{n} |d_\nu|^2 k_\nu,$$

as[15] for $\mu \neq \nu$

$$\text{(29a)} \qquad \int_{-1}^{1} l_\mu(t) l_\nu(t) p_1(t)\, dt = \frac{1}{\omega_n'(x_\mu)\omega_n'(x_\nu)} \int_{-1}^{1} \frac{\omega_n(x)}{(x - x_\mu)(x - x_\nu)}\, \omega_n(x) p_1(x)\, dx = 0,$$

and

$$\text{(29b)} \qquad \int_{-1}^{1} l_\mu(x)^2 p_1(x)\, dx = \int_{-1}^{1} l_\mu(x) p_1(x)\, dx = k_\mu.$$

As

$$F(x_0) = 1 = \sum_{\nu=1}^{n} d_\nu l_\nu(x_0),$$

we obtain from what precedes

$$1 = \left| \sum_{\nu=1}^{n} d_\nu l_\nu(x_0) \right|^2 = \left| \sum_{\nu=1}^{n} d_\nu \sqrt{k_\nu}\, \frac{l_\nu(x_0)}{\sqrt{k_\nu}} \right|^2 \leqq \left(\sum_{\nu=1}^{n} \frac{|l_\nu(x_0)|^2}{k_\nu} \right) \left(\sum_{\nu=1}^{n} k_\nu |d_\nu|^2 \right),$$

i.e.

$$I(F) \geqq \frac{1}{\displaystyle\sum_{\nu=1}^{n} \frac{|l_\nu(x_0)|^2}{k_\nu}}.^{16}$$

[15] P. Erdös and P. Turán: *On Interpolation. I.* Annals of Math., 1937, pp. 142–155.

[16] Implicitly J. Shohat: *Théorie générale etc.*, p. 47, formula (75).

Equality is evidently to be obtained if and only if

$$F(x) = \frac{1}{\sum_{\nu=1}^{n} \frac{|l_\nu(x_0)|^2}{k_\nu}} \cdot \sum_{\nu=1}^{n} \frac{\overline{l_\nu(x_0)}}{k_\nu} l_\nu(x),$$

i.e.

$$(30) \qquad \sum_{\nu=1}^{n} \frac{|l_\nu(x_0)|^2}{k_\nu} = \left[\min_{F(x_0)=1,\ F(x)=a_0+\cdots+a_{n-1}x^{n-1}} \int_{-1}^{1} |F(t)|^2 p_1(t)\, dt \right]^{-1}.$$

But then evidently

$$(31) \qquad \begin{aligned} \sum_{\nu=1}^{n} \frac{|l_\nu(x_0)|^2}{k_\nu} &\leqq \left[\min_{F(x_0)=1,\ F(x)=a_0+\cdots+a_{n-1}x^{n-1}} \int_{-1}^{1} |F(t)|^2 p_2(t)\, dt \right]^{-1} \\ &= \sum_{\nu=1}^{n} \frac{|l_\nu^+(x_0)|^2}{k_\nu^+} \qquad \text{Q.e.d.} \end{aligned}$$

In the special case of x_0 being the ν^{th} root of the n^{th} polynomial orthogonal to $p_1(x)$, then by (30) the minimum-value is k_ν and this minimum is attained only for $F(x) = l_\nu(x)$ (Corollary I).[17]

Here we remark—although we make no use of it in this paper—that the sum $\sum_{\nu=1}^{n} \frac{l_\nu^{(r)}(x)^2}{k_\nu}$ is also monotone with respect to $p_1(x)$, if $l_\nu^{(r)}(x)$ denotes the r^{th} derivative.

In Lemma II let $p_1(x) \equiv p(x)$ and $p_2(x) = m(>0)$ if $a \leqq x \leqq b$, and $p_2(x) = 0$ for the complementary intervals; furthermore suppose x_0 real. Then the explicit form of the polynomials orthogonal in $[-1, +1]$ with respect to $p_2(x)$ is given by

$$\omega_n(x) = A \cdot P_n\left(-1 + 2\frac{x-a}{b-a}\right),$$

where $P_n(x)$ denotes the n^{th} Legendre-polynomial for $[-1, +1]$, with the normalization $P_n(1) = 1$, A depending only upon n, a, b so that coeff. x^n in $\omega_n(x)$ equals 1. As in this case with $P_n(\eta_\nu) = 0$ $(\nu = 1, 2, \cdots n)$,

$$k_\nu^+ = m \int_a^b l_\nu^+(t)\, dt = \frac{m(b-a)}{2} \int_{-1}^{1} \frac{P_n(t)}{P_n'(\eta_\nu)(t-\eta_\nu)}\, dt = \frac{m(b-a)}{(1-\eta_\nu^2)P_n'(\eta_\nu)^2},\text{[18]}$$

from Lemma II, if $\xi_1, \xi_2, \cdots, \xi_n$ denote the roots of $P_n\left(-1 + 2\frac{x-a}{b-a}\right) = 0$,

$$(32) \qquad \begin{aligned} \sum_{\nu=1}^{n} \frac{l_\nu(x_0)^2}{k_\nu} &\leqq \frac{1}{m(b-a)} \sum_{\nu=1}^{n} (1-\eta_\nu^2)P_n'(\eta_\nu)^2 \frac{P_n\left(-1+2\frac{x_0-a}{b-a}\right)^2}{\left(\frac{2}{b-a}\right)^2 P_n'(\eta_\nu)^2 (x_0-\xi_\nu)^2} \\ &= \frac{b-a}{4m} P_n\left(-1+2\frac{x_0-a}{b-a}\right)^2 \sum_{\nu=1}^{n} \frac{1-\eta_\nu^2}{(x_0-\xi_\nu)^2}. \end{aligned}$$

[17] J. Shohat: *On the convergence-properties of Lagrange-interpolation etc.*, Annals of Math., 1937, pp. 758–769, formula (39).

[18] L. Fejér: *Az interpolátióról*, Akadémiai Értesitö (Hungarian), 1915.

As $-1 + 2\dfrac{\xi_\nu - a}{b-a} = \eta_\nu$ ($\nu = 1, 2, \cdots n$), making use of the differential equation of the Legendre-polynomials we obtain by the substitution $x_0 = a + \dfrac{b-a}{2}(1 + y_0)$ in (32) for any real y_0

$$(33)\quad \sum_{\nu=1}^{n} \frac{l_\nu\left(a + \frac{b-a}{2}(y_0+1)\right)^2}{k_\nu} \leqq \frac{1}{m(b-a)}[(1-y_0^2)P_n'(y_0)^2 + n^2P_n(y_0)^2].$$

Equality in (33) holds only for $p(x) \equiv m$. Suppose now $-1 \leqq y_0 \leqq +1$; then, by a well known result $|P_n(y_0)| \leqq 1$, and by the above cited theorem of Bernstein-Fejér $|P_n'(y_0)\sqrt{(1-y_0^2)}| \leqq n$, i.e. from (33) for $a \leqq x \leqq b$ we obtain

$$(34a)\quad \sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} < \frac{2}{m(b-a)}n^2.$$[19]

Let now $-1 + \epsilon \leqq y \leqq 1 - \epsilon$. Then according to the classical formula of Laplace for $\epsilon' \leqq \vartheta \leqq \pi - \epsilon'$, we have

$$\left| P_n(\cos\vartheta) - \left[\frac{2}{\pi n \sin\vartheta}\right]^{\frac{1}{2}} \cos\left[\left(n + \frac{1}{2}\right)\vartheta - \frac{\pi}{4}\right]\right| < \frac{c_{31}(\epsilon')}{n^{3/2}};$$

by this and by the theorem of Bernstein-Fejér for $[-1 + \epsilon, 1 - \epsilon]$ we have

$$|P_n(x)| \leqq \frac{c_{32}(\epsilon)}{\sqrt{n}}, \qquad |P_n'(x)| \leqq c_{32}(\epsilon)\sqrt{n},$$

i.e. we obtain roughly[19] from (33) in $[a + \epsilon, b - \epsilon]$

$$(34b)\quad \sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} \leqq \frac{c_{33}(\epsilon)}{m(b-a)}n.$$

We remark, that for the validity of (34a) and (34b) in the above intervals we require only that the L-integrable $p(x)$ is in $[-1, +1]$ not less than 0, and in $[a, b]$ $p(x) \geqq m > 0$. (Corollary II.)

Let in Lemma II $p_1(x) \equiv p(x)$, $p_2(x) = 0$ in $[-1, a][b, 1]$ and not less than $\dfrac{m}{[(x-a)(b-x)]^{\frac{1}{2}}}$ in $[a, b]$; if $\gamma_1, \gamma_2, \cdots \gamma_n$ stand for the roots of the Tchebycheff polynomial $T_n(x)(T_n(\cos\vartheta) = \cos n\vartheta)$, $\mu_1, \mu_2, \cdots \mu_n$ for the roots of $T_n\left(-1 + \dfrac{2(x-a)}{b-a}\right)$, then

$$k_\nu^+ = m\int_a^b \frac{b_\nu^+(t)}{[(t-a)(b-t)]^{\frac{1}{2}}}dt = m\int_{-1}^{1} \frac{T_n(t)}{T_n'(\gamma_\nu)(t-\gamma_\nu)}\frac{dt}{\sqrt{(1-t^2)}} = m\frac{\pi}{n},$$

[19] For $a = -1$, $b = 1$, see J. Shohat: *On Interpolation*, Annals of Math., 1933.

i.e. by easy computation with real x

$$\sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} \leqq \frac{n}{m\pi} \sum_{\nu=1}^{n} l_\nu^+(x)^2$$

(34c)

$$= \frac{n}{m\pi}\left[1 - \frac{1}{2n} + \frac{1}{2n(2n-1)} T'_{2n-1}\left(-1 + \frac{2(x-a)}{b-a}\right)\right].$$

Thus we came to the result, that if the L-integrable $p(x)$ is not less in $[-1, +1]$ than 0 and in the subinterval $[a, b]$ $p(x) \geqq \dfrac{m}{[(x-a)(b-x)]^{\frac{1}{2}}}$, then for $a \leqq x \leqq b$

(35)
$$\sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} \leqq \frac{2}{\pi m} \cdot n.$$

(Corollary III). Equality holds only when in $[a, b]$ $p(x) = \dfrac{m}{[(x-a)(b-x)]^{\frac{1}{2}}}$, and in the complementary intervals $p(x) = 0$, further $x = a$ or $x = b$.

We deduce theorem II from (34a) and (34b) as follows. As

$$\sum_{\nu=1}^{n} k_\nu^{(n)} = \int_{-1}^{1}\left(\sum_{\nu=1}^{n} l_\nu(t)\right) p(t)\, dt = \int_{-1}^{1} p(t)\, dt$$

and as $k_\nu^{(n)} > 0$ by (6a) and (29b), we have in $[a, b]$ by (34a)

$$\left(\sum_{\nu=1}^{n} | l_\nu(x) |\right)^2 = \left(\sum_{\nu=1}^{n} \frac{| l_\nu(x) |}{\sqrt{k_\nu}} \sqrt{k_\nu}\right)^2 \leqq \sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} \cdot \sum_{\nu=1}^{n} k_\nu < \frac{2}{m(b-a)} \int_{-1}^{1} p(t)\, dt \cdot n^2$$

and analogously in $[a + \epsilon, b - \epsilon]$

$$\left(\sum_{\nu=1}^{n} | l_\nu(x) |\right)^2 < \frac{c_{33}(\epsilon)}{m(b-a)} \int_{-1}^{1} p(t)\, dt \cdot n.$$

But then

$$|\omega_n(x)|\, 2^{n-2} \leqq 2 \sum_{\nu=1}^{n} |l_\nu(x)| \begin{cases} \leqq \left[\dfrac{8}{m(b-a)}\right]^{\frac{1}{2}}\left[\displaystyle\int_{-1}^{1} p(t)\, dt\right]^{\frac{1}{2}} \cdot n & \text{for } a \leqq x \leqq b \\ \leqq \left[\dfrac{4c_{33}(\epsilon)}{m(b-a)}\right]^{\frac{1}{2}}\left[\displaystyle\int_{-1}^{1} p(t)\, dt\right]^{\frac{1}{2}} \cdot \sqrt{n} & \text{for } a + \epsilon \leqq x \leqq b - \epsilon. \end{cases}$$

It is to be remarked, that the same argument leads to the following result: if the L-integrable $p(x)$ is not less than 0 in $[-1, +1]$ and if in the subinterval $[a, b]$

$$p(x) \geqq \frac{m}{[(x-a)(b-x)]^{\frac{1}{2}}},$$

then for $a \leqq x \leqq b$ we have

(36)
$$| \omega_n(x) | \leqq 4\left[\frac{2}{\pi m} \int_{-1}^{1} p(t)\, dt\right]^{\frac{1}{2}} \cdot \frac{\sqrt{n}}{2^n} \qquad n = 1, 2, \cdots.$$

Some further corollaries of Lemma II we shall mention later.

As $-1 + 2\dfrac{\xi_\nu - a}{b - a} = \eta_\nu$ $(\nu = 1, 2, \cdots n)$, making use of the differential equation of the Legendre-polynomials we obtain by the substitution $x_0 = a + \dfrac{b-a}{2}(1 + y_0)$ in (32) for any real y_0

$$(33)\quad \sum_{\nu=1}^{n} \frac{l_\nu\left(a + \frac{b-a}{2}(y_0 + 1)\right)^2}{k_\nu} \leqq \frac{1}{m(b-a)}\,[(1 - y_0^2)P_n'(y_0)^2 + n^2 P_n(y_0)^2].$$

Equality in (33) holds only for $p(x) \equiv m$. Suppose now $-1 \leqq y_0 \leqq +1$; then, by a well known result $|P_n(y_0)| \leqq 1$, and by the above cited theorem of Bernstein-Fejér $|P_n'(y_0)\sqrt{(1 - y_0^2)}| \leqq n$, i.e. from (33) for $a \leqq x \leqq b$ we obtain

$$(34a)\quad \sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} < \frac{2}{m(b-a)}\, n^2.^{19}$$

Let now $-1 + \epsilon \leqq y \leqq 1 - \epsilon$. Then according to the classical formula of Laplace for $\epsilon' \leqq \vartheta \leqq \pi - \epsilon'$, we have

$$\left| P_n(\cos\vartheta) - \left[\frac{2}{\pi n \sin\vartheta}\right]^{\frac{1}{2}} \cos\left[\left(n + \frac{1}{2}\right)\vartheta - \frac{\pi}{4}\right]\right| < \frac{c_{31}(\epsilon')}{n^{3/2}};$$

by this and by the theorem of Bernstein-Fejér for $[-1 + \epsilon, 1 - \epsilon]$ we have

$$|P_n(x)| \leqq \frac{c_{32}(\epsilon)}{\sqrt{n}}, \qquad |P_n'(x)| \leqq c_{32}(\epsilon)\sqrt{n},$$

i.e. we obtain roughly[19] from (33) in $[a + \epsilon, b - \epsilon]$

$$(34b)\quad \sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} \leqq \frac{c_{33}(\epsilon)}{m(b-a)}\, n.$$

We remark, that for the validity of (34a) and (34b) in the above intervals we require only that the L-integrable $p(x)$ is in $[-1, +1]$ not less than 0, and in $[a, b]$ $p(x) \geqq m > 0$. (Corollary II.)

Let in Lemma II $p_1(x) \equiv p(x)$, $p_2(x) = 0$ in $[-1, a][b, 1]$ and not less than $\dfrac{m}{[(x-a)(b-x)]^{\frac{1}{2}}}$ in $[a, b]$; if $\gamma_1, \gamma_2, \cdots \gamma_n$ stand for the roots of the Tchebycheff polynomial $T_n(x)(T_n\ (\cos\vartheta) = \cos n\vartheta)$, $\mu_1, \mu_2, \cdots \mu_n$ for the roots of $T_n\left(-1 + \dfrac{2(x-a)}{b-a}\right)$, then

$$k_\nu^+ = m\int_a^b \frac{b_\nu^+(t)}{[(t-a)(b-t)]^{\frac{1}{2}}}\, dt = m\int_{-1}^{1} \frac{T_n(t)}{T_n'(\gamma_\nu)(t - \gamma_\nu)}\, \frac{dt}{\sqrt{(1-t^2)}} = m\,\frac{\pi}{n},$$

[19] For $a = -1$, $b = 1$, see J. Shohat: *On Interpolation*, Annals of Math., 1933.

i.e. by easy computation with real x

$$\sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} \leqq \frac{n}{m\pi} \sum_{\nu=1}^{n} l_\nu^{+}(x)^2 \tag{34c}$$
$$= \frac{n}{m\pi}\left[1 - \frac{1}{2n} + \frac{1}{2n(2n-1)}\, T'_{2n-1}\left(-1 + \frac{2(x-a)}{b-a}\right)\right].$$

Thus we came to the result, that if the L-integrable $p(x)$ is not less in $[-1, +1]$ than 0 and in the subinterval $[a, b]$ $p(x) \geqq \dfrac{m}{[(x-a)(b-x)]^{\frac{1}{2}}}$, then for $a \leqq x \leqq b$

$$\sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} \leqq \frac{2}{\pi m} \cdot n. \tag{35}$$

(Corollary III). Equality holds only when in $[a, b]$ $p(x) = \dfrac{m}{[(x-a)(b-x)]^{\frac{1}{2}}}$, and in the complementary intervals $p(x) = 0$, further $x = a$ or $x = b$.

We deduce theorem II from (34a) and (34b) as follows. As

$$\sum_{\nu=1}^{n} k_\nu^{(n)} = \int_{-1}^{1} \left(\sum_{\nu=1}^{n} l_\nu(t)\right) p(t)\, dt = \int_{-1}^{1} p(t)\, dt$$

and as $k_\nu^{(n)} > 0$ by (6a) and (29b), we have in $[a, b]$ by (34a)

$$\left(\sum_{\nu=1}^{n} |\, l_\nu(x)\, |\right)^2 = \left(\sum_{\nu=1}^{n} \frac{|\, l_\nu(x)\, |}{\sqrt{k_\nu}} \sqrt{k_\nu}\right)^2 \leqq \sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} \cdot \sum_{\nu=1}^{n} k_\nu < \frac{2}{m(b-a)} \int_{-1}^{1} p(t)\, dt \cdot n^2$$

and analogously in $[a + \epsilon, b - \epsilon]$

$$\left(\sum_{\nu=1}^{n} |\, l_\nu(x)\, |\right)^2 < \frac{c_{33}(\epsilon)}{m(b-a)} \int_{-1}^{1} p(t)\, dt \cdot n.$$

But then

$$|\omega_n(x)|\, 2^{n-2} \leqq 2 \sum_{\nu=1}^{n} |l_\nu(x)| \begin{cases} \leqq \left[\dfrac{8}{m(b-a)}\right]^{\frac{1}{2}} \left[\displaystyle\int_{-1}^{1} p(t)\, dt\right]^{\frac{1}{2}} \cdot n & \text{for } a \leqq x \leqq b \\[2ex] \leqq \left[\dfrac{4c_{33}(\epsilon)}{m(b-a)}\right]^{\frac{1}{2}} \left[\displaystyle\int_{-1}^{1} p(t)\, dt\right]^{\frac{1}{2}} \cdot \sqrt{n} & \text{for } a + \epsilon \leqq x \leqq b - \epsilon. \end{cases}$$

It is to be remarked, that the same argument leads to the following result: if the L-integrable $p(x)$ is not less than 0 in $[-1, +1]$ and if in the subinterval $[a, b]$

$$p(x) \geqq \frac{m}{[(x-a)(b-x)]^{\frac{1}{2}}},$$

then for $a \leqq x \leqq b$ we have

$$|\, \omega_n(x)\, | \leqq 4\left[\frac{2}{\pi m} \int_{-1}^{1} p(t)\, dt\right]^{\frac{1}{2}} \cdot \frac{\sqrt{n}}{2^n} \qquad n = 1, 2, \cdots. \tag{36}$$

Some further corollaries of Lemma II we shall mention later.

Let us now consider the lower estimate of the orthogonal polynomials $\omega_n(x)$.

THEOREM III. *Let the weight-function $p(x)$ be non-negative and L-integrable in $[-1, +1]$; throughout the subinterval $[a, b]$ suppose $p(x) \geqq m > 0$: Then, if $x_d^{(n)}$ denotes* the root of $\omega_n(x)$ *nearest to x, we have for real x*

$$|\omega_n(x)| \geqq \left[\frac{c_{34} m}{(b-a)\int_{-1}^{1} p(t)\,dt}\right]^{\frac{1}{2}} \left(\frac{b-a}{4}\right)^n |x - x_d^{(n)}|.$$

We require two lemmas.

LEMMA III. *In $[-1, +1]$ suppose $p_1(x) \geqq p_2(x) \geqq 0$ and both L-integrable. If $\omega_n(x)$ and $\omega_n^+(x)$ denote the corresponding orthogonal polynomials respectively, k_ν and k_ν^+ the respective Christoffel-numbers, x_ν and x_ν^+ the respective fundamental points, then*

$$\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)^2} \leqq \sum_{\nu=1}^{n} \frac{1}{k_\nu^+ \omega_n^{+\prime}(x_\nu^+)^2}.$$

PROOF. Let us consider the minimum of $N(F) = \int_{-1}^{1} |F(t)|^2 p_1(t)\,dt$ amongst the polynomials of degree $(n-1)$, in which coeff. $x^{n-1} = 1$. It is known that this problem has one and only one solution and that the minimum is assumed only for $F(x) = \omega_{n-1}(x)$. But we want to represent the solution in the form $F(x) = \sum_{\nu=1}^{n} d_\nu l_\nu(x)$, where the $l_\nu(x)$'s denote the fundamental functions of Lagrange-interpolation formed upon the roots of $\omega_n(x)$. It is evident that in this case we have to determine the minimum of the form $\sum_{\nu=1}^{n} k_\nu |d_\nu|^2$, if $\sum_{\nu=1}^{n} \frac{d_\nu}{\omega_n'(x_\nu)} = 1$. From this, by applying Schwarz's inequality once more, we obtain for the value of this minimum $\left[\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)^2}\right]^{-1}$ and equality holds only for the polynomial

$$f(x) = \frac{\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)} l_\nu(x)}{\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)^2}}. \tag{37}$$

Thus again

$$\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)^2} = \left[\min_{f=x^{n-1}+\cdots} \int_{-1}^{1} |f(t)|^2 p_1(t)\,dt\right]^{-1} \leqq \left[\min_{f=x^{n-1}+\cdots} \int_{-1}^{1} |f(t)|^2 p_2(t)\,dt\right]^{-1}$$

$$= \sum_{\nu=1}^{n} \frac{1}{k_\nu^+ \omega_n^{+\prime}(x_\nu)^2}. \qquad \text{Q.e.d.}$$

If $p_1(x) \equiv p(x)$ and throughout the intervals $[-1, a]$, $[b, 1]$ $p_2(x) = 0$, further for $a \leqq x \leqq b$ $p_2(x) \equiv m$, then

$$\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)^2} \leqq \frac{1}{m}\left[\min_{f=x^{n-1}+\cdots} \int_a^b f(t)^2\,dt\right]^{-1} = \frac{1}{m}\left[\int_a^b P_{n-1}\left(-1 + 2\frac{t-a}{b-a}\right)^2 dt\right]^{-1},$$

where the integrand is the linear transform of the $(n-1)^{\text{th}}$ Legendre polynomial with the normalization coeff. $x^{n-1}=1$. Thus

$$\sum_{\nu=1}^{n}\frac{1}{k_\nu\omega_n'(x_\nu)^2}\leqq\frac{1}{m}\frac{(2n-1)\binom{2n-2}{n-1}^2}{(b-a)^{2n-1}},\tag{38}$$

if throughout $[-1,+1]$ the L-integrable $p(x)\geqq 0$ and throughout $[a,b]$ $p(x)\geqq m>0$. Equality holds only if $p(x)\equiv m$ in $[a,b]$ and $\equiv 0$ outside of $[a,b]$.

LEMMA IV. *For the fundamental functions of the Lagrange-interpolation formed upon any matrix* $\mathfrak{M}$ *we have*

$$\phi(x)=l_{kn}(x)+l_{k+1,n}(x)\geqq 1.$$
$$x_k^{(n)}\geqq x\geqq x_{k+1}^{(n)}.$$

PROOF. Let $2<k<n-2$. Then $\phi(x)$ is a polynomial of degree $(n-1)$ at the utmost, vanishing at $x_1^{(n)}, x_2^{(n)}\cdots, x_{k-1}^{(n)}, x_{k+2}^{(n)},\cdots x_n^{(n)}$, i.e. at $n-2$ places and equals 1 at $x_k^{(n)}$ and $x_{k+1}^{(n)}$. Consequently its first derivative has one root in each of the intervals $[x_2^{(n)}, x_1^{(n)}]\cdots[x_{k-1}^{(n)}, x_{k-2}^{(n)}]$, $[x_{k+3}^{(n)}, x_{k+2}^{(n)}],\cdots[x_n^{(n)}, x_{n-1}^{(n)}]$, which determines at least $n-4$ roots. In consequence of $\phi(x_k^{(n)})=\phi(x_{k+1}^{(n)})=1$ one of the roots of this derivative lies evidently in $[x_{k+1}^{(n)}, x_k^{(n)}]$. We now show that $\phi'(x_k^{(n)})\leqq 0$ and $\phi'(x_{k+1}^{(n)})\geqq 0$. It will be sufficient to show the first. Suppose $\phi'(x_k)>0$. Then $\phi'(x)$ must have at least one root in $[x_k^{(n)}, x_{k-1}^{(n)}]$. But then $\phi'(x)$ could not have more roots and thus $\phi'(x_{k+1})<0$; hence $\phi'(x)$ ought to have one more root in $[x_{k+2}^{(n)}, x_{k+1}^{(n)}]$, and this is impossible. $\phi'(x_{k+1}^{(n)})\geqq 0$ is to be obtained analogously. But then, if in $(x_{k+1}^{(n)}, x_k^{(n)})$ there were a point ξ_0 with $\phi(\xi_0)=1$ then $\phi'(x)$ would have 3 roots in $[x_{k+1}^{(n)}, x_k^{(n)}]$; an evident impossibility, which establishes the lemma. For $k=1, 2, n-2, n-1$ the proof runs analogously.

PROOF OF THEOREM III. Here also we start from $\sum_{\nu=1}^{n}\frac{l_\nu(x)^2}{k_\nu}$. By (38) we have—$x_d^{(n)}$ has the meaning given above—

$$\begin{aligned}\omega_n(x)^2&=\left[\sum_{\mu=1}^{n}\frac{1}{k_\mu\omega_n'(x_\mu)^2(x-x_\mu)^2}\right]^{-1}\sum_{\mu=1}^{n}\frac{l_\mu(x)^2}{k_\mu}\\&\geqq|x-x_d|^2\frac{m(b-a)^{2n-1}}{(2n-1)\binom{2n-2}{n-1}^2}\sum_{\nu=1}^{n}\frac{l_\nu(x)^2}{k_\nu}.\end{aligned}\tag{39}$$

But since $k_\nu^{(n)}<\int_{-1}^{1}p(t)\,dt$ (we made use of this at the upper estimate, too) then if e.g. $x_{d+1}^{(n)}\leqq x\leqq x_d^{(n)}$, we may write

$$\omega_n(x)^2>\frac{m}{\int_{-1}^{1}p(t)\,dt}\frac{(b-a)^{2n-1}}{(2n-1)\binom{2n-2}{n-1}^2}[l_d(x)^2+l_{d+1}(x)^2]\,|x-x_d^{(n)}|^2.\tag{40}$$

Now by lemma IV for any x in the interval in question

$$l_d(x)^2 + l_{d+1}(x)^2 \geqq \tfrac{1}{2}[l_d(x) + l_{d+1}(x)]^2 \geqq \tfrac{1}{2}$$

i.e. by

$$(2n-1)\binom{2n-2}{n-1}^2 \sim c_{35} 4^n$$

we obtain that

$$|\,\omega_n(x)\,| > c_{36}\left\{\frac{m}{b-a}\left[\int_{-1}^{1} p(t)\,dt\right]^{-1}\right\}^{\frac{1}{2}}\left(\frac{b-a}{4}\right)^n |\,x - x_d^{(n)}\,|,$$

which proves the theorem.

Theorem IV. *Let us add to the hypotheses of theorem III that, throughout a subinterval* $[c, d]$ *of* $[a, b]$, $m \leqq p(x) \leqq \dfrac{M}{\sqrt{(1-x^2)}}$. *Then, if* $x_d^{(n)}$ *denotes again the root of* $\omega_n(x)$ *nearest to* x, *we have in* $[c + \epsilon, d - \epsilon]$ *for* $n > n_0(\epsilon, c, d, p)$

$$|\,\omega_n(x)\,| > \frac{c_{37}}{\sqrt{(b-a)}}\left[\frac{m}{M + \int_{-1}^{1} p(t)\,dt}\right]^{\frac{1}{2}} |\,x - x_d^{(n)}\,|\left(\frac{b-a}{4}\right)^n \sqrt{n}.$$

For the proof we require

Lemma V. *In* $[-1, +1]$ *suppose* $p(x) \geqq 0$ *and L-integrable; suppose furthermore, throughout a subinterval* $[u, v]$ $p(x) \leqq \dfrac{M}{\sqrt{(1-x^2)}}$; *then for the Christoffel-numbers* $k_\nu^{(n)}$ *belonging to the* $x_\nu^{(n)}$*'s lying in* $[u + \eta, v - \eta]$ $(\eta > 0)$, *we have*

$$k_\nu^{(n)} < \frac{c_{37}}{n}\left[M + \frac{c_{39}}{\eta^2 n}\int_{-1}^{1} p(t)\,dt\right].$$

Proof. According to the first corollary due to Shohat of Lemma II we may write

$$k_\nu^{(n)} = \min_{f(x_\nu^{(n)})=1,\, f=a_0x^{n-1}+\cdots} \int_{-1}^{1} |f(t)|^2 p(t)\,dt$$

$$= \min_{F(\vartheta_\nu^{(n)})=1,\, F(\vartheta)=b_0+\cdots+b_{n-1}\cos(n-1)\vartheta} \int_0^{\pi} |F(\vartheta)|^2 p(\cos\vartheta)\sin\vartheta\,d\vartheta$$

$$\leqq \int_0^{\pi}\left[\left(\frac{\sin n\dfrac{\vartheta+\vartheta_\nu^{(n)}}{2}}{n\sin\dfrac{\vartheta+\vartheta_\nu^{(n)}}{2}}\right)^2 + \left(\frac{\sin n\dfrac{\vartheta-\vartheta_\nu^{(n)}}{2}}{n\sin\dfrac{\vartheta-\vartheta_\nu^{(n)}}{2}}\right)^2\right] p(\cos\vartheta)\sin\vartheta\,d\vartheta$$

$$\equiv \int_\alpha^\beta + \int_C \equiv I_1 + I_2,$$

where $u = \cos\beta$, $v = \cos\alpha$ and C stands for the intervals complementary to $[\alpha, \beta]$ in $[0, \pi]$. Evidently

$$I_1 < M \frac{2}{n^2} \int_0^{2\pi} \left(\frac{\sin n \dfrac{\vartheta - \vartheta_\nu^{(n)}}{2}}{\sin \dfrac{\vartheta - \vartheta_\nu^{(n)}}{2}} \right)^2 d\vartheta = \frac{c_{38} M}{n}$$

and

$$I_2 < \frac{c_{39}}{\eta^2 n^2} \int_{-1}^{1} p(t)\, dt,$$

which proves the lemma.

PROOF OF THEOREM IV. From (39)

$$\omega_n(x)^2 \geqq m \frac{(b-a)^{2n-1}(x - x_d^{(n)})^2}{(2n-1)\binom{2n-2}{n-1}^2} \left[\frac{l_d(x)^2}{k_d} + \frac{l_{d+1}(x)^2}{k_{d+1}} \right].$$

If $c + \epsilon \leqq x \leqq d - \epsilon$ and $n > n_1(c, d, \epsilon, p)$ then, according to footnote,[7] the interval $[x_{d+1}^{(n)}, x_d^{(n)}]$ containing x lies in $[c + \frac{1}{2}\epsilon, d - \frac{1}{2}\epsilon]$, and hence, by applying lemma IV, and with the substitution $u = c$, $v = d$, $\eta = \frac{1}{2}\epsilon$, Lemma V we obtain for $n > \max(n_1, 4/\epsilon^2) = n_0$

$$\omega(x)^2 > c_{40} \frac{mn}{M + \int_{-1}^{1} p(t)\, dt} \left(\frac{b-a}{4}\right)^{2n-1} | x - x_d^{(n)} |^2,$$

which establishes the theorem.

REMARK I. In the special case $m \leqq p(x) \leqq \dfrac{M}{\sqrt{(1-x^2)}}$ throughout $[-1, +1]$ we have by the aforesaid in $[-1 + \epsilon, 1 - \epsilon]$

$$c_{41}(p, \epsilon) \frac{\sqrt{n}}{2^n} | x - x_d^{(n)} | \leqq | \omega_n(x) | \leqq c_{42}(p, \epsilon) \frac{\sqrt{n}}{2^n}.$$

REMARK II. In lemma V we required the upper estimate of the Christoffel-numbers. Although we shall not use it, we mention, that if in $[-1, +1]$ $p(x) \geqq m > 0$, then

$$k_\nu^{(n)} \geqq \frac{2m}{(1 - x_\nu^2) P_n'(x_\nu)^2 + n^2 P_n(x_\nu)^2},$$

equality only for $p(x) \equiv m$; here $P_n(x)$ means the n^{th} Legendre-polynomial with normalization $P_n(1) = 1$. By this

$$k_\nu^{(n)} \geqq \frac{m}{n^2}, \qquad \nu = 1, 2, \cdots n, \qquad n = 1, 2, \cdots.$$

Further it is easy to obtain that if in $[-1, +1]$ $p(x) \geqq m > 0$, then

$$k_\nu^{(n)} \geqq \frac{c_{43}(\epsilon)m}{n}, \qquad -1 + \epsilon \leqq x_\nu^{(n)} \leqq 1 - \epsilon,$$

and, if in $[-1, +1]$ $p(x) \geqq \dfrac{m}{\sqrt{(1 - x^2)}}$ holds, then

$$k_\nu^{(n)} \geqq \frac{c_{44}m}{n}, \qquad \nu = 1, 2, \cdots n.$$

REMARK III. If in Lemma V $[u, v] \equiv [-1, +1]$, then we have the sharper result

$$k_\nu^{(n)} \leqq \frac{c_{45}M}{n}, \qquad \nu = 1, 2, \cdots n.$$

REMARK IV. We obtain from the proof of Lemma III

$$\omega_{n-1}(x) = \frac{\displaystyle\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)} l_\nu(x)}{\displaystyle\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)^2}} = \frac{\displaystyle\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)^2} \cdot \frac{\omega_n(x)}{x - x_\nu}}{\displaystyle\sum_{\nu=1}^{n} \frac{1}{k_\nu \omega_n'(x_\nu)^2}},$$

where $l_\nu(x)$ are the fundamental functions of the Lagrange interpolation formed upon the roots of $\omega_n(x)$. As $k_\nu > 0$ we evidently have for sufficiently small ϵ (see (2))

$$\operatorname{sg} \omega_{n-1}(x_\nu^{(n)} - \epsilon) = -\operatorname{sg} \omega_n(x_\nu^{(n)} - \epsilon),$$
$$\operatorname{sg} \omega_{n-1}(x_{\nu+1}^{(n)} + \epsilon) = \operatorname{sg} \omega_n(x_{\nu+1}^{(n)} + \epsilon),$$

i.e. we obtained the well known fact that there is a root of $\omega_{n-1}(x)$ between each pair of roots of $\omega_n(x)$.

2.

THEOREM V. *If for the fundamental functions belonging to the matrix* $\mathfrak{M}$

(41) $$[| l_k(x) |]^{1/n} \leqq 1 + \epsilon, \qquad -1 \leqq x \leqq +1, k = 1, 2, \cdots n, n > n_2(\epsilon),$$

holds for any sufficiently small ϵ, *then, at any fixed point* z *of the complex plane cut up along* $[-1, +1]$, *we have*

$$\lim_{n\to\infty} [\omega_n(z)]^{1/n} = \frac{z + \sqrt{(z^2 - 1)}}{2}.$$

Here we must take those values of the roots which are positive on the positive real axis for $z > 1$.

For the proof we require

LEMMA VI. *From* (41) *it follows that, for any small positive* η, *if* $n > n_3(\eta)$,

$$| \omega_n'(x_\nu) | > (\tfrac{1}{2} - \eta)^n, \qquad \nu = 1, 2, \cdots n.$$

PROOF. Suppose that the lemma is not true. In this case there is a positive absolute constant δ and a sequence of positive integers $n_1 < n_2 < \cdots$ such, that to any n_k we could give an integer ν_k with $1 \leqq \nu_k \leqq n_k$ and

$$|\omega'_{n_k}(x_{\nu_k}^{(n_k)})| < (\tfrac{1}{2} - \delta)^{n_k}.$$

But, as, according to a classical theorem of Tchebycheff, there is in $[-1, +1]$ a ξ_k for which the value of the polynomial $\dfrac{\omega_{n_k}(x)}{x - x_{\nu_k}^{(n_k)}}$ is not less than $\dfrac{1}{2^{n_k-2}}$, at the same $x = \xi_k$

$$|l_{\nu_k, n_k}(\xi_k)| > \frac{1}{2^{n_k-2}} \cdot \frac{1}{(\frac{1}{2} - \delta)^{n_k}}$$

which, for $k \to \infty$ contradicts (41) and thus proves the lemma.

REMARK. It follows from lemma VI and lemma I that (41) implies

$$\lim_{n\to\infty} \left[\sum_{\nu=1}^{n} \frac{1}{|\omega'_n(x_\nu)|} \right]^{1/n} = 2.$$

We shall make no use of this statement in this paper.

PROOF OF THEOREM V. Let

$$\eta_\nu = \cos \frac{2\nu - 1}{2n + 2}\pi, \qquad \nu = 1, 2, \cdots (n + 1),$$

the roots of the $(n + 1)^{\text{th}}$ Tchebycheff-polynomial $T_{n+1}(x)$ ($T_n(\cos \vartheta) = \cos n\vartheta$), and represent $\omega_n(x)$ by the Lagrange-interpolatoric-polynomial taken at these η_ν's. If $L_\nu(z)$ ($\nu = 1, 2 \cdots (n + 1)$) denote *these* fundamental functions, then

$$\omega_n(z) = \sum_{\nu=1}^{n+1} \omega_n(\eta_\nu) L_\nu(z) \tag{42}$$

(z any number). But by (41) and lemma I, in $-1 \leqq x \leqq +1$, for $n > n_2(\epsilon)$ we have

$$|\omega_n(x)| 2^{n-2} \leqq 2 \sum_{\nu=1}^{n} \left| \frac{\omega_n(x)}{\omega'_n(x_\nu)(x - x_\nu)} \right| \leqq 2(n + 1)(1 + \epsilon)^n$$

i.e. by $|T'_{n+1}(\eta_\nu)| = \dfrac{n + 1}{\sqrt{(1 - \eta_\nu^2)}}$ from (42) for $n > n_3(\epsilon)$

$$|\omega_n(z)| < \sum_{\nu=1}^{n+1} \frac{8(n + 1)(1 + \epsilon)^n}{2^n} \frac{|z + \sqrt{(z^2 - 1)}|^{n+1}}{(n + 1)|z - \eta_\nu|}$$

$$< 17 \left| \frac{1 + \epsilon}{2} (z + \sqrt{(z^2 - 1)}) \right|^{n+1} \max_{\nu=1,2,\cdots(n+1)} \frac{1}{|z - \eta_\nu|};$$

hence, as z is not in $[-1, +1]$, for $n > n_4(\epsilon, z)$

$$[|\omega_n(z)|]^{1/n} < \frac{1 + 2\epsilon}{2} |z + \sqrt{(z^2 - 1)}|.$$

Now let us consider the lower estimate and represent the $(n-1)^{\text{th}}$ Tchebycheff-polynomial $T_{n-1}(z)$ by the Lagrange-interpolatoric-polynomial taken at the roots of $\omega_n(x)$. We obtain

$$T_{n-1}(z) \equiv \sum_{\nu=1}^{n} T_{n-1}(x_\nu^{(n)}) l_\nu(x).$$

As $|T_{n-1}(x_\nu)| \leqq 1$, there is an integer ν_0 with $1 \leqq \nu_0 \leqq n$ and

$$|l_{\nu_0}(z)| > \frac{1}{n} |T_{n-1}(z)|,$$

thus for $n > n_6(\epsilon)$

$$(43) \qquad |\omega_n(z)| > \frac{1}{2n}(1-\epsilon)|z + \sqrt{(z^2-1)}|^{n-1} |\omega_n'(x_{\nu_0}^{(n)})| |z - x_{\nu_0}^{(n)}|.$$

From (43), by lemma VI, and as z is not at $[-1, +1]$, we obtain for $n > n_6(\epsilon)$

$$(44) \qquad |\omega_n(z)| \geqq \frac{1}{2n} \left| \frac{1-2\epsilon}{2}(z + \sqrt{(z^2-1)}) \right|^{n-1} |z - x_{\nu_0}^{(n)}|.$$

From (44) and $n > n_7(\epsilon, B)$ it is evident that in any closed bounded set B of the complex plane cut up along $[-1, +1]$ we have uniformly

$$(45) \qquad [|\omega_n(z)|]^{1/n} > (1-3\epsilon) \left| \frac{z + \sqrt{(z^2-1)}}{2} \right|,$$

where we are to take that value of the square root, for which the right side $\sim z$ for $z \to \infty$. This proves the theorem, since on the positive real axis, for $z \to \infty$, the two sides of (45) are equal without the sign of the absolute value too, and upon the cut plane both sides are one-valued and regular functions of z.

We already mentioned in the introduction that in the case of strongly normal polynomials, the asymptotic formula for $[\omega_n(z)]^{1/n}$ upon the cut plane is a consequence of the above Fejérian-theorem. In order to state an analogous theorem for a class of orthogonal polynomials more general than that of Szegö's we require a further lemma.

Lemma VII. *If the L-integrable weight-function $p(x)$ is non-negative in $[-1, +1]$ and if its roots form an aggregate of measure 0, then for the fundamental functions connected with the matrix p and $n > n_8(\epsilon)$ we have*

$$|l_{\nu,n}(x)| \leqq (1+\epsilon)^n, \qquad -1 \leqq x \leqq +1, \qquad \nu = 1, 2, \cdots, n.$$

Proof. We employ following theorem of E. Remes.[20] Let in $[-1, +1]$ be given a finite set of disjoint intervals of total length ϑ. If throughout this

[20] E. Remes: *Sur une propriété extrêmale des polynomes de Tchebichef*. Communications de l'Institut de Sciences etc., Kharkow, 1936, série 4, XIII, fasc. 1, pp. 93-95.

aggregate the absolute value of the polynomial $f(x)$ of degree n is not greater than M then in $[-1, +1]$

$$|f(x)| \leqq M \left| T_n\left(\frac{4}{\vartheta} - 1\right) \right|,$$

where $T_n(\cos\vartheta) = \cos n\vartheta$.

Now suppose lemma VII to be untrue. Then we have an infinite sequence of positive integers $n_1 < n_2 < \cdots$ and a c_{46} such, that to every n_k there is a positive integer ν_k with $1 \leqq \nu_k \leqq n_k$ and a ξ_{ν_k} with $-1 \leqq \xi_{\nu_k} \leqq +1$ to satisfy

$$(46) \qquad |l_{\nu_k, n_k}(\xi_k)| > (1 + c_{46})^{n_k}.$$

Apply now for $l_{\nu_k, n_k}(x)$ Remes's theorem taking for the aggregate of intervals of length ϑ those intervals throughout which

$$|l_{\nu_k, n_k}(x)| \leqq \left(1 + \frac{c_{46}}{2}\right)^{n_k}.$$

Thus we obtain

$$\left(1 + \frac{c_{46}}{2}\right)^{n_k} \left| T_{n_k}\left(\frac{4}{\vartheta} - 1\right) \right| \geqq (1 + c_{46})^{n_k}$$

from which $0 < \vartheta < 2 - c_{47}$ where c_{47} depends only upon c_{46} and is independent of n_k. Hence throughout intervals $\mathfrak{A}$ of total length greater than c_{47}, $|l_{\nu_k, n_k}(x)| \geqq (1 + \frac{1}{2}c_{46})^{n_k}$. But then by the assumption made for the roots of $p(x)$ we may omit from $\mathfrak{A}$ intervals of smaller total length than $\frac{1}{2}c_{47}$ such that throughout the remaining $\mathfrak{A}^+$ $p(x) \geqq c_{48}$, where c_{48} depends only upon c_{47}. Hence we had

$$\int_{-1}^{1} l_{\nu_k, n_k}(x)^2 p(x)\, dx \geqq \int_{\mathfrak{A}+} l_{\nu_k, n_k}(x)^2 p(x)\, dx$$

$$\geqq c_{48} \int_{\mathfrak{A}+} l_{\nu_k, n_k}(x)^2\, dx > c_{48} \cdot \frac{c_{47}}{2} \left(1 + \frac{c_{46}}{2}\right)^{n_k}$$

which contradicts the Shohat minimum property of the $l_\nu(x)$'s (Lemma II. Coroll. I.). Hence lemma VII is established.

By theorem V and lemma VII we state

THEOREM VI. *Let the weight-function $p(x)$ be non-negative and L-integrable in $[-1, +1]$ and assume the aggregate of its roots to be of measure 0. Then, taking the suitable values of the roots we have upon the plane cut up along $[-1, +1]$*

$$\lim_{n \to \infty} [\omega_n(z)]^{1/n} = \frac{z + \sqrt{(z^2 - 1)}}{2}$$

uniformly in each interior domain.

3.

In this section we consider c) problems. We prove following Fejérian theorem:

THEOREM VII. *Assume the matrix $\mathfrak{M}$ to be such that* $[-1, +1]$ *possesses a subinterval* $[b, a] \equiv [\cos\beta, \cos\alpha]$ *with*

$$| l_k(x) | \leqq c_{49}, \qquad k = \nu, \nu + 1, \cdots \mu$$

and for the other fundamental functions

$$| l_k(x) | < c_{50} n^{c_{51}}$$

throughout $[b, a]$, *if*

$$\vartheta_{\nu-1}^{(n)} < \alpha \leqq \vartheta_\nu^{(n)} < \vartheta_{\nu+1}^{(n)} < \cdots < \vartheta_\mu^{(n)} \leqq \beta < \vartheta_{\mu+1}^{(n)}.$$

Then

$$\frac{[\epsilon(b - a)]^{\frac{1}{2}}}{c_{49}} \cdot \frac{1}{n} \leqq \vartheta_{k+1}^{(n)} - \vartheta_k^{(n)} \leqq \frac{c_{49} \cdot c_{52}(\epsilon, a, b, c_{50}, c_{51})}{n},$$

if $\vartheta_k^{(n)}$ *and* $\vartheta_{k+1}^{(n)}$ *are in* $[\alpha + \epsilon, \beta - \epsilon]$, ϵ *denoting any small positive number.*

PROOF. First we prove the lower estimate

$$\frac{1}{| \vartheta_{k+1}^{(n)} - \vartheta_k^{(n)} |} = \left| \frac{l_k(\cos\vartheta_k^{(n)}) - l_k(\cos\vartheta_{k+1}^{(n)})}{\vartheta_k^{(n)} - \vartheta_{k+1}^{(n)}} \right| = \left| \frac{dl_k(\cos\vartheta)}{d\vartheta} \right|_{\vartheta=\vartheta^+},$$

where $\vartheta_k^{(n)} \leqq \vartheta^+ \leqq \vartheta_{k+1}^{(n)}$. But then, $l_k(\cos\vartheta)$ being a trigonometric polynomial of order $(n - 1)$, by the Bernstein-Fejér-theorem we have

$$\frac{1}{| \vartheta_{k+1}^{(n)} - \vartheta_k^{(n)} |} \leqq \frac{c_{49}(n - 1)}{[\epsilon(b - a)]^{\frac{1}{2}}},$$

which proves the lower estimate.

Let us now consider the upper estimate. Suppose

$$\max(\vartheta_{l+1}^{(n)} - \vartheta_l^{(n)}) = \frac{2A(n)}{n} = \vartheta_{k+1}^{(n)} - \vartheta_k^{(n)}, \qquad \alpha + \epsilon \leqq \vartheta_l^{(n)} < \vartheta_{l+1}^{(n)} \leqq \beta - \epsilon$$

and we have to prove that $A(n)$ remains smaller than a number independent of n. We can suppose $A(n) > 10$. Let r be the smallest positive even integer greater than $[c_{51}] + 4$, μ the largest integer with $\frac{\mu - 1}{2} r \leqq n - 1$, $\frac{\vartheta_k^{(n)} + \vartheta_{k+1}^{(n)}}{2} = \delta_0$ and

$$\varphi(\vartheta) = \frac{1}{\mu^r}\left[\left(\frac{\sin\mu\frac{\vartheta + \delta_0}{2}}{\sin\frac{\vartheta + \delta_0}{2}}\right)^r + \left(\frac{\sin\mu\frac{\vartheta - \delta_0}{2}}{\sin\frac{\vartheta - \delta_0}{2}}\right)^r\right]. \tag{47}$$

Then $\varphi(\vartheta)$ is a non-negative, pure cosine polynomial of order $(n - 1)$ at the utmost, for which

$$\varphi(\delta_0) \geqq 1, \tag{48a}$$

and if, without any loss of generality, we assume $0 \leqq \delta_0 \leqq \frac{\pi}{2}$,

$$| \varphi(\vartheta) | \leqq \frac{1}{\mu^r}\left[\frac{1}{\sin^r \dfrac{\vartheta + \delta_0}{2}} + \frac{1}{\sin^r \dfrac{\vartheta - \delta_0}{2}}\right] \leqq \left(\frac{9\pi^2}{2\mu}\right)^r \frac{2}{(\vartheta - \delta_0)^r}. \tag{48b}$$

Let us now represent $\varphi(\vartheta)$ by the n^{th} Lagrange-interpolatory polynomial taken upon $\mathfrak{M}$. Then by (48a) and (48b) we have

$$\begin{aligned} 1 \leqq | \varphi(\delta_0) | &= \left| \sum_{\nu=1}^{n} \varphi(\vartheta_\nu^{(n)}) l_\nu(\cos \delta_0) \right| < \sum_{\nu=1}^{n} \left(\frac{9\pi^2}{2\mu}\right)^r \frac{| l_\nu(\cos \delta_0) |}{(\vartheta_\nu^{(n)} - \delta_0)^r} \\ &< c_{49} \sum_{\alpha + \frac{\epsilon}{2} \leqq \vartheta_\nu^{(n)} \leqq \beta - \frac{\epsilon}{2}} \frac{1}{(\vartheta_\nu^{(n)} - \delta_0)^r} \cdot \left(\frac{9\pi^2}{2\mu}\right)^r + \left(\frac{9\pi^2}{2\mu}\right)^r c_{50} n^{c_{51}} \cdot \left(\frac{2}{\epsilon}\right)^r n. \end{aligned} \tag{49}$$

Let us divide the sum upon the right-hand side into two parts with $\nu \leqq k$ or $\nu \geqq k + 1$ respectively. As, according to the already proved lower estimate in the first sum we have

$$\delta_0 - \vartheta_\nu^{(n)} > \frac{A(n)}{n} + \frac{(k - \nu)[\epsilon(b - a)]^{\frac{1}{2}}}{\sqrt{2}\, c_{49} n} \qquad \text{for} \quad \vartheta_\nu^{(n)} \geqq \alpha + \frac{\epsilon}{2}$$

and in the second one

$$\vartheta_\nu^{(n)} - \delta_0 > \frac{A(n)}{n} + \frac{(k + 1 - \nu)[\epsilon(b - a)]^{\frac{1}{2}}}{\sqrt{2}\, c_{49} n} \qquad \text{for} \quad \vartheta_\nu^{(n)} \leqq \beta - \frac{\epsilon}{2},$$

we obtain that the sum on the right hand side of (49)

$$\begin{aligned} &< 2n^r \sum_{\nu=0}^{\infty} \frac{1}{\left(A(n) + \nu \dfrac{[\epsilon(b - a)]^{\frac{1}{2}}}{c_{49}\sqrt{2}}\right)^r} \\ &\qquad < \frac{2n^r}{r - 1} \frac{c_{49}\sqrt{2}}{[\epsilon(b - a)]^{\frac{1}{2}}} \frac{1}{\left(A - \dfrac{[\epsilon(b - a)]^{\frac{1}{2}}}{c_{49}\sqrt{2}}\right)^{r-1}} < \frac{c_{53}(\epsilon, a, b) c_{49} n^r}{\left(\dfrac{A}{2}\right)^{r-1}}, \end{aligned}$$

as $r \geqq 4$ and $\frac{A}{2} > 5 > \frac{[\epsilon(b - a)]^{\frac{1}{2}}}{c_{49}\sqrt{2}}$ $(c_{49} \geqq 1)$. By substituting this into (49) we obtain

$$1 < \frac{c_{54}(c_{51}, \epsilon, c_{50})}{n^2} + \frac{c_{49}^2 c_{55}(\epsilon, a, b, c_{51})}{A^2},$$

which establishes the upper estimate.

The consequences of this theorem for sequences of strongly normal polynomials, we already mentioned in the introduction. For orthogonal polynomials we have

THEOREM VIII. *If the weight-function $p(x)$ is non-negative and L-integrable throughout* $[-1, +1]$, *and if, throughout the subinterval* $[b, a] \equiv [\cos \beta, \cos \alpha]$ $0 < m \leqq p(x) \leqq \dfrac{M}{\sqrt{(1-x^2)}}$, *then for any* $\epsilon > 0$

$$\frac{c_{56}(a, b, p, \epsilon)}{n} \leqq \vartheta_{k+1}^{(n)} - \vartheta_k^{(n)} \leqq \frac{c_{57}(a, b, p, \epsilon)}{n}$$

if $\alpha + \epsilon \leqq \vartheta_k^{(n)} < \vartheta_{k+1}^{(n)} \leqq \beta - \epsilon$.

PROOF. By (34a) we have for $b \leqq x \leqq a$

$$\sum_{\nu=1}^{n} \frac{l_\nu(x)^2}{k_\nu} < \frac{2}{m(a-b)} n^2.$$

As $k_\nu < \int_{-1}^{1} p(t)\, dt$, for $b \leqq x \leqq a$,

$$(50) \quad |l_\nu(x)| < \left[\frac{2}{m(a-b)} \int_{-1}^{1} p(t)\, dt\right]^{\frac{1}{2}} \cdot n, \qquad \nu = 1, 2, \cdots n, \qquad n = 1, 2, \cdots.$$

Further *a fortiori* from (34b)

$$\frac{l_\mu(x)^2}{k_\mu} \leqq \frac{c_{33}\left(\frac{\epsilon}{2}\right)}{m(a-b)} n, \qquad \mu = 1, 2, \cdots n, \qquad b + \frac{\epsilon}{2} \leqq x \leqq a - \frac{\epsilon}{2},$$

i.e. by applying lemma V with $u = b$, $v = a$, $\eta = \frac{1}{2}\epsilon$ for those $l_\mu(x)$, for which $b + \frac{1}{2}\epsilon \leqq x_\mu^{(n)} \leqq a - \frac{1}{2}\epsilon$ we obtain

$$|l_\mu(x)|^2 < \frac{c_{33}\left(\frac{\epsilon}{2}\right)}{m(a-b)} n \cdot \frac{c_{58}(p, \epsilon)}{n} = \frac{c_{59}(p, \epsilon)}{a-b},$$

$$b + \frac{\epsilon}{2} \leqq x \leqq a - \frac{\epsilon}{2}.$$

Thus the premises of theorem VII are satisfied for $[b + \frac{1}{2}\epsilon, a - \frac{1}{2}\epsilon]$ with $c_{49} = \left[\dfrac{c_{59}(p, \epsilon)}{a-b}\right]^{\frac{1}{2}}$, $c_{50} = \left[\dfrac{2}{m(a-b)} \displaystyle\int_{-1}^{1} p(t)\, dt\right]^{\frac{1}{2}}$, $c_{51} = 1$; hence theorem VIII is proved.

REMARK I. From corollary III of lemma II and from remark III of theorem IV it follows that if $\dfrac{m}{\sqrt{(1-x^2)}} \leqq p(x) \leqq \dfrac{M}{\sqrt{(1-x^2)}}$ in $[-1, +1]$, then

$$\sum_{\nu=1}^{n} l_\nu(x)^2 \leqq c_{60} \frac{M}{m}.$$

REMARK II. By theorems IV and VIII we obtain that if in $[-1, +1]$ the non-negative $p(x)$ is L-integrable and in a subinterval $[b, a]$ $0 < m \leqq p(x) \leqq$

$\frac{M}{\sqrt{(1-x^2)}}$, then $|\omega_n(x)|$ takes between any two roots lying in $[b+\epsilon, a-\epsilon]$ a value greater than $\frac{c_{61}(a, b, p, \epsilon)}{\sqrt{n}}\left(\frac{b-a}{4}\right)^n$.

We already mentioned in the introduction that if in the subinterval the weight-function is supposed to be continuous there are asymptotic theorems to be obtained. For sake of simplicity let us ascribe to $p(x)$ besides the properties of being non-negative and L-integrable in $[-1, +1]$ also continuity and positiveness of $p(x)\sqrt{(1-x^2)}$ throughout $[-1, +1]$. The results are new for this case, too, but the argument is very much clearer.

In (34c) for $x = \xi_0 = \cos\varphi_0$ we had

$$\min_{\substack{f(t)=c_0+\cdots+c_{n-1}t^{n-1}\\ f(\xi_0)=1}} \int_{-1}^{1} \frac{f(t)^2}{\sqrt{(1-t^2)}}\,dt = \frac{\pi}{n-\frac{1}{2}+\frac{1}{2}\frac{\sin(2n-1)\varphi_0}{\sin\varphi_0}}. \tag{51}$$

Consider the polynomial

$$\phi_{n-1}(x) = \frac{T_{n-1}(\xi_0)T_n(x) - T_n(\xi_0)T_{n-1}(x)}{x-\xi_0}\cdot\frac{1}{n-\frac{1}{2}+\frac{1}{2}\frac{\sin(2n-1)\varphi_0}{\sin\varphi_0}} \tag{52}$$

with $T_0(x) = 1/\sqrt{2}$, $T_\nu(\cos\vartheta) = \cos\nu\vartheta$, $\nu \geqq 1$. It is evident that for $n > 1$ $\phi_{n-1}(\xi_0) = 1$ and that by the formula of Christoffel-Darboux (which holds for $n > 1$)

$$\frac{T_{n-1}(\xi_0)T_n(x) - T_n(\xi_0)T_{n-1}(x)}{x-\xi_0} = 2\sum_{\nu=0}^{n-1} T_\nu(\xi_0)T_\nu(x) \tag{53}$$

we may write for $n > 1$

$$\phi_{n-1}(x) \equiv \frac{\sum_{\nu=0}^{n-1} T_\nu(\xi_0)T_\nu(x)}{\sum_{\nu=0}^{n-1} T_\nu(\xi_0)^2} \equiv 2\frac{\sum_{\nu=0}^{n-1} T_\nu(\xi_0)T_\nu(x)}{n-\frac{1}{2}+\frac{1}{2}\frac{\sin(2n-1)\varphi_0}{\sin\varphi_0}}, \tag{54}$$

which means that $\phi_{n-1}(x)$ is that polynomial which, amongst the $f(x)$ polynomials of degree $(n-1)$ with $f(\xi_0) = 1$ minimizes the integral $\int_{-1}^{1} f(t)^2 \frac{dt}{\sqrt{(1-t^2)}}$. By (54) we have for any $-1 \leqq \xi_0 \leqq 1$, $-1 \leqq x \leqq +1$

$$|\phi_{n-1}(x)| \leqq c_{61}. \tag{55}$$

For $\phi_{n-1}(x)$ we require following two lemmas.

Lemma VIII. *If for $n \to \infty$ $n\varphi_0 \to +\infty$ and $n(\pi - \varphi_0) \to \infty$, then the distance between φ_0 and the root of $\phi_{n-1}(\cos\vartheta) = 0$ nearest to $\vartheta = \varphi_0$ is $\sim\pi/n$.*

Proof. According to (52) it will suffice to consider the roots differing from $\vartheta = \varphi_0$ of the equation

$$\cos(n-1)\varphi_0\cos n\vartheta - \cos n\varphi_0\cdot\cos(n-1)\vartheta = 0. \tag{56}$$

As each of the $(n-1)$ intervals given by the n different real roots of $T_n(x) = 0$ contains just one root of the equation $T_{n-1}(x) = 0$ it is clear that any equation of the form $\lambda T_n(x) + \mu T_{n-1}(x)$ hence also $\phi_{n-1}(x) = 0$, have a root in the interval $\left[\cos \frac{2l+1}{2n}\pi, \cos \frac{2l-1}{2n}\pi\right]$ $(l = 1, \cdots (n-1))$. Thus equation (56) has a root in each interval of length $2\pi/n$ and consequently for the rightwards root φ_0' next to φ_0 we surely have $|\varphi_0' - \varphi_0| \leqq 2\pi/n$. Then by simple transformation we obtain from (56)

$$\sin \frac{\vartheta - \varphi_0}{2} \sin (n - \tfrac{1}{2})(\vartheta + \varphi_0) - \sin \frac{\vartheta + \varphi_0}{2} \cdot \sin (n - \tfrac{1}{2})(\varphi_0 - \vartheta) = 0$$

which for $n\varphi_0 \to \infty$, $n(\pi - \varphi_0) \to \infty$ immediately leads to $A \to \pi$ if $\varphi_0' = \varphi_0 + A/n$.

Lemma IX. *If for the weight-function $p(x)$ in $[-1, +1]$, $q(x) \equiv p(x)\sqrt{(1-x^2)} \geqq m > 0$ and $q(x)$ is continuous throughout the same interval, then for $0 \leqq \varphi_0 \leqq \pi$, $\xi_0 = \cos \varphi_0$*

$$\lim_{n\to\infty} n \int_{-1}^{1} \phi_{n-1}(t)^2 p(t)\, dt = \pi p(\xi_0)\sqrt{(1-\xi_0^2)} \lim_{n\to\infty} \frac{1}{1 + \frac{1}{2n} \frac{\sin (2n-1)\varphi_0}{\sin \varphi_0}}$$

i.e. if $n\varphi_0 \to \infty$, $n(\pi - \varphi_0) \to \infty$

$$\lim_{n\to\infty} n \int_{-1}^{1} \phi_{n-1}(t)^2 p(t)\, dt = \pi p(\xi_0)\sqrt{(1-\xi_0^2)}.$$

Proof. For $|x - \xi_0| \geqq 1/n^{\frac{1}{4}}$ it is easy to obtain from (52)

$$|\phi_{n-1}(x)| \leqq 20n^{-\frac{1}{4}}$$

i.e.

$$(57) \qquad \int\limits_{\substack{-1\leqq t\leqq +1\\ |t-\xi_0|\geqq n^{-1/4}}} \phi_{n-1}(t)^2 p(t)\, dt < \frac{400}{n^{3/2}} \int_{-1}^{1} p(t)\, dt.$$

If $|p(x')\sqrt{(1-x'^2)} - p(\xi_0)\sqrt{(1-\xi_0^2)}| \leqq \delta$ for any $|x' - \xi_0| \leqq n^{-\frac{1}{4}}$, we have

$$(58) \qquad \left| \int\limits_{\substack{-1\leqq t\leqq +1\\ |t-\xi_0|\leqq n^{-1/4}}} \phi_{n-1}(t)^2 p(t)\, dt - p(\xi_0)\sqrt{(1-\xi_0^2)} \int_{-1}^{1} \frac{\phi_{n-1}(t)^2}{\sqrt{(1-t^2)}}\, dt \right|$$

$$< \frac{400}{n^{3/2}}\pi \max_{-1\leqq x\leqq +1} p(x)\sqrt{(1-x^2)} + \delta \int_{-1}^{1} \frac{\phi_{n-1}(t)^2}{\sqrt{(1-t^2)}}\, dt.$$

From (57) and (58) evidently

$$\left|\int_{-1}^{1}\phi_{n-1}(t)^2 p(t)\,dt - \frac{\pi p(\xi_0)\sqrt{(1-\xi_0^2)}}{n-\dfrac{1}{2}+\dfrac{1}{2}\dfrac{\sin(2n-1)\varphi_0}{\sin\varphi_0}}\right|$$

$$= \left|\int_{-1}^{1}\phi_{n-1}(t)^2\left[p(t) - \frac{p(\xi_0)\sqrt{(1-\xi_0^2)}}{\sqrt{(1-t^2)}}\right]dt\right|$$

$$< \frac{400\pi}{n^{3/2}}\left[\int_{-1}^{1}p(t)\,dt + \max_{-1\le x\le +1} p(x)\sqrt{(1-x^2)}\right] + \frac{10\pi\delta}{n}. \quad \text{Q.e.d.}$$

REMARK. If $p(x)\sqrt{(1-x^2)} = q(x)$ has at $x = \xi_0$ a discontinuity of first order, then $q(\xi_0)$ at the right-hand side in Lemma IX is replaced by $\dfrac{q(\xi_0+0)+q(\xi_0-0)}{2}$.

LEMMA X. *If for the polynomial $f(x)$ of degree k $f(\xi_0) = 1$, $\xi_0 = \cos\varphi_0$ and any positive ϵ and $\delta < 1$, $k \geqq 16/\delta$ and $\sin\varphi_0 > 8/k\delta$ the inequality*

$$\int\limits_{\substack{|t-\xi_0|\le\epsilon\\ -1\le t\le +1}} \frac{f(t)^2}{\sqrt{(1-t^2)}}\,dt \leqq (1-\delta)\frac{\pi}{k+\dfrac{1}{2}+\dfrac{1}{2}\dfrac{\sin(2k+1)\varphi_0}{\sin\varphi_0}}$$

holds, then

$$\int_{-1}^{1}\frac{f(t)^2}{\sqrt{(1-t^2)}}\,dt > \frac{\delta\left(1+\dfrac{\epsilon^2}{4}\right)^{\frac{1}{4}k\delta}}{400k}.$$

REMARK. This lemma means that if the quadratic integral of the polynomial normalized for 1 at $x = \xi_0$ is "too small" in the interval $[\xi_0 - \epsilon, \xi_0 + \epsilon]$, it must be very large in some other parts of $[-1, +1]$.

PROOF. Without any loss of generality we may suppose $\xi_0 \geqq 0$. Then construct with the above $f(x)$

$$F(x) = f(x)\left[1 - \left(\frac{x-\xi_0}{1+\xi_0}\right)^2\right]^{[\frac{1}{4}k\delta]} \tag{59}$$

a polynomial, whose degree is less than $k(1 + \frac{1}{2}\delta)$. In $[-1, +1]$ we evidently have $\left|\dfrac{x-\xi_0}{1+\xi_0}\right| \leqq 1$. From (34c)

$$\int_{-1}^{1}\frac{F(t)^2}{\sqrt{(1-t^2)}}\,dt \geqq \frac{\pi}{k+2\left[\dfrac{k\delta}{4}\right]+\dfrac{1}{2}+\dfrac{1}{2}\dfrac{\sin\left(2k+4\left[\dfrac{k\delta}{4}\right]+1\right)\varphi_0}{\sin\varphi_0}}; \tag{60}$$

on the other hand

$$\int_{-1}^{1}\frac{F(t)^2}{\sqrt{(1-t^2)}}\,dt < \left(1-\frac{\epsilon^2}{4}\right)^{[\frac{1}{4}k\delta]}\int_{|t-\xi_0|\geqq\epsilon}\frac{f(t)^2}{\sqrt{(1-t^2)}}\,dt + \int_{\xi_0-\epsilon}^{\xi_0+\epsilon}\frac{f(t)^2}{\sqrt{(1-t^2)}}\,dt, \tag{61}$$

as in $[-1, +1]$ the second factor of $F(x)$ is non-negative but $\leqq 1$. Making use of the hypothesis after the arrangement we obtain from (60) and (61)

$$\int_{-1}^{1} \frac{f(t)^2}{\sqrt{(1-t^2)}}\,dt > \int_{|t-\xi_0|\geqq\epsilon} \frac{f(t)^2}{\sqrt{(1-t^2)}}\,dt$$

$$> \frac{\pi}{\left(1-\frac{\epsilon^2}{4}\right)^{\frac{1}{4}k\delta}}\left[\frac{1}{k+\frac{k\delta}{2}+\frac{1}{2}+\frac{1}{2}\frac{\sin\left(2k+4\left[\frac{k\delta}{4}\right]+1\right)\varphi_0}{\sin\varphi_0}}\right.$$

$$\left.-\frac{1-\delta}{k+\frac{1}{2}+\frac{1}{2}\frac{\sin(2k+1)\varphi_0}{\sin\varphi_0}}\right] > \frac{\left(1+\frac{\epsilon^2}{4}\right)^{\frac{1}{4}k\delta}}{100k^2}\left[\frac{k\delta}{2}-\frac{2}{\sin\varphi_0}\right] > \frac{\delta\left(1+\frac{\epsilon^2}{4}\right)^{\frac{1}{4}k\delta}}{400k}.$$

Q.e.d.

According to what has been said before we may deduce asymptotic formulas for the Christoffel numbers belonging to $p(x)$.

THEOREM IX. *Let $p(x)\sqrt{(1-x^2)}$ be continuous and $p(x)\sqrt{(1-x^2)} \geqq m > 0$ in $[-1, +1]$; then, if $n \to \infty$, we have for any $x_\nu^{(n)}$ lying in* $-\left[1-\frac{\log n}{n^2}\right]^{\frac{1}{2}} \leqq x_\nu^{(n)} \leqq \left[1-\frac{\log n}{n^2}\right]^{\frac{1}{2}}$

$$k_\nu^{(n)} = \int_{-1}^{1} l_\nu(t)^2 p(t)\,dt \sim \frac{\pi p(x_\nu^{(n)})\sqrt{(1-x_\nu^{(n)2})}}{n}.$$

PROOF. First we show that we have for any ϵ and δ independent of n if $n > n_8(\delta, \epsilon)$ and $|x_\nu^{(n)}| \leqq \left[1-\frac{64}{n^2\delta^2}\right]^{\frac{1}{2}}$

$$\int\limits_{\substack{-1\leqq t\leqq +1\\ |t-x_\nu^{(n)}|\leqq\epsilon}} \frac{l_\nu(t)^2}{\sqrt{(1-t^2)}}\,dt > (1-\delta)\frac{\pi}{n-\frac{1}{2}+\frac{1}{2}\frac{\sin(2n-1)\vartheta_\nu^{(n)}}{\sin\vartheta_\nu^{(n)}}}. \tag{62}$$

Suppose the contrary, then by lemma X we had

$$\int_{-1}^{1} \frac{l_\nu(t)^2}{\sqrt{(1-t^2)}}\,dt > \frac{\delta\left(1+\frac{\epsilon^2}{4}\right)^{\frac{1}{4}\delta n}}{500n},$$

i.e. *a fortiori*

$$\int_{-1}^{1} l_\nu(t)^2 p(t)\,dt \geqq m\,\frac{\delta\left(1+\frac{\epsilon^2}{4}\right)^{\frac{1}{4}\delta n}}{500n},$$

which contradicts the minimum-property of Shohat if n is sufficiently large. Thus (62) is proved. But then by (62) for $n > n_0(\delta, \epsilon)$

$$\int_{-1}^{1} l_\nu(t)^2 p(t)\,dt > \min_{|x-x_\nu^{(n)}|\leq\epsilon} p(x)\sqrt{(1-x^2)} \int_{\substack{|x-x_\nu^{(n)}|\leq\epsilon \\ -1\leq x\leq+1}} \frac{l_\nu(t)^2}{\sqrt{(1-t^2)}}\,dt$$

$$\tag{63} > (1-\delta)\,\frac{\pi p(x_\nu^{(n)})\sqrt{(1-x_\nu^{(n)2})}}{n-\dfrac{1}{2}+\dfrac{1}{2}\dfrac{\sin(2n-1)\vartheta_\nu^{(n)}}{\sin\vartheta_\nu^{(n)}}}.$$

On the other hand by the minimum-property

$$\tag{64} \int_{-1}^{1} l_\nu(t)^2 p(t)\,dt \leqq \int_{-1}^{1} \phi_{n-1}(t)^2 p(t)\,dt \sim \frac{\pi p(x_\nu^{(n)})\sqrt{(1-x_\nu^{(n)2})}}{n-\dfrac{1}{2}+\dfrac{1}{2}\dfrac{\sin(2n-1)\vartheta_\nu^{(n)}}{\sin\vartheta_\nu^{(n)}}}$$

by lemma IX, if we replace ξ_0 by $x_\nu^{(n)}$; (63) and (64) lead to theorem IX.

And now we may go over to the asymptotic representation of the fundamental functions.

THEOREM X. *In* $[-1, +1]$ *let* $p(x)\sqrt{(1-x^2)}$ *be continuous and such that* $p(x)\sqrt{(1-x^2)} \geqq m > 0$. *Then, if* $\epsilon > 0$, $n > n_0(\epsilon)$ $|x_\nu^{(n)}| \leqq \left[1 - \frac{\log n}{n^2}\right]^{\frac{1}{2}}$, *for* $-1 \leqq x \leqq +1$ *we have uniformly* ($T_r(\cos\vartheta) = \cos r\vartheta$, $r > 1$)

$$|l_{\nu,n}(x) - \phi_{n-1}(x)| \equiv \left| l_\nu(x) - \frac{T_{n-1}(x_\nu^{(n)})T_n(x) - T_n(x_\nu^{(n)})T_{n-1}(x)}{\left(n-\dfrac{1}{2}+\dfrac{1}{2}\dfrac{\sin(2n-1)\vartheta_\nu^{(n)}}{\sin\vartheta_\nu^{(n)}}\right)(x-x_\nu^{(n)})}\right| < \epsilon.$$

PROOF. Let us consider, with the above $\phi_{n-1}(x)$, the integral

$$\tag{65} I_\nu \equiv \int_{-1}^{1} [l_\nu(t) - \phi_{n-1}(t)]^2 p(t)\,dt.$$

As $\phi_{n-1}(x) \equiv \sum_{k=1}^{n} \phi(x_k^{(n)}) l_k(x)$ and $\phi_{n-1}(x_\nu^{(n)}) = 1$, we obtain

$$\tag{66} \begin{aligned} I_\nu &= k_\nu^{(n)} - 2\sum_{d=1}^{n} \phi_{n-1}(x_d^{(n)}) \int_{-1}^{1} l_d(t) l_\nu(t) p(t)\,dt \\ &\quad + \int_{-1}^{1} \phi_{n-1}(t)^2 p(t)\,dt = -k_\nu^{(n)} + \int_{-1}^{1} \phi_{n-1}(t)^2 p(t)\,dt. \end{aligned}$$

But then by theorem IX and lemma IX, for $n \to \infty$, $|x_\nu^{(n)}| \leqq \left[1 - \frac{\log n}{n^2}\right]^{\frac{1}{2}}$ uniformly in ν

$$\tag{67} \lim_{n\to\infty} nI_\nu = 0.$$

From the premise and the remark to theorem VIII it follows for $[-1, +1]$ and $\nu = 1, 2, \cdots n$

$$| l_\nu(x) | \leqq c_{60} \frac{M}{m}.$$

By this and (55) for $l_\nu (\cos \vartheta) - \phi_{n-1} (\cos \vartheta) = \psi(\vartheta)$ we obtain

$$| \psi(\vartheta) | \leqq c_{62}(p).$$

But then by the theorem of Bernstein-Fejér, in $[0, \pi]$ we have

$$(68) \qquad | \psi'(\vartheta) | \leqq c_{62}(p)n.$$

Let $\psi(\vartheta)$ assume its absolute maximum in $[0, \pi]$ at $\vartheta = \gamma_0$ and let this maximum value be D_ν. Then in the subinterval $i \equiv \left[\gamma_0 - \dfrac{D_\nu}{2c_{62}n}, \gamma_0 + \dfrac{D_\nu}{2c_{62}n}\right]$ of $[0, \pi]$, by (68) we have

$$|\psi(\vartheta)| > D_\nu - \frac{D_\nu}{2} = \frac{D_\nu}{2}$$

i.e.

$$I_\nu \equiv \int_{-1}^{1} [l_\nu(t) - \phi_{n-1}(t)]^2 p(t)\, dt = \int_0^\pi \psi(\vartheta)^2 p\, (\cos \vartheta) \sin \vartheta\, d\vartheta$$

$$> m \int_{(i)} \psi(\vartheta)^2\, d\vartheta > m \frac{D_\nu}{c_{62}n} \cdot \frac{D_\nu^2}{4}.$$

Thus by (67), for $n \to \infty$ *a fortiori*

$$\frac{mD_\nu^3}{4c_{62}} \to 0$$

i.e. $D_\nu \to 0$, which establishes the theorem.

From theorem IX and lemma VIII we easily deduce

THEOREM XI. *Let $p(x)\sqrt{(1 - x^2)}$ be continuous in $[-1, +1]$ with $p(x)\sqrt{(1 - x^2)} \geqq m > 0$, further let $C(n)$, for $n \to \infty$ arbitrarily slowly, tend to infinity; then for those roots* $\cos \vartheta_\nu^{(n)}$ *of the* n^{th} *polynomial orthogonal to $p(x)$, which satisfy*

$$\frac{C(n)}{n} \leqq \vartheta_k^{(n)} < \vartheta_{k+1}^{(n)} \leqq \pi - \frac{C(n)}{n},$$

we have uniformly in k

$$\lim_{n \to \infty} n(\vartheta_{k+1}^{(n)} - \vartheta_k^{(n)}) = \pi.$$

REMARK I. If we assume not $p(x)\sqrt{(1 - x^2)}$, but $p(x)$ to be $\geqq m$ and continuous and not in $[-1, +1]$ but in the subinterval $[a, b]$, then $\phi_{n-1}(x)$ is to be replaced by a polynomial of similar form but the Tchebycheff polynomials

$T_n(x)$ are to be replaced by $P_n(x)$ Legendre-polynomials. Theorem IX and X remain true for those k_ν and $l_\nu(x)$, for which $a + \epsilon \leqq x_\nu^{(n)} \leqq b - \epsilon$ and $a + \epsilon \leqq x \leqq b - \epsilon$.

Remark II. If we attribute again continuity and positiveness in $[-1, +1]$ to $p(x)\sqrt{(1 - x^2)}$, it is probable, that theorem X holds for the fundamental functions belonging to *every* $x_\nu^{(n)}$. Theorem XI does not hold for *every* $x_\nu^{(n)}$ in the original form; the difference $\vartheta_{k+1}^{(n)} - \vartheta_k^{(n)}$ will be asymptotically equal to the distance between $\vartheta_k^{(n)}$ and that root of

$$\cos (n - 1)\vartheta_k^{(n)} \cos n\vartheta - \cos n\vartheta_k^{(n)} \cos (n - 1)\vartheta = 0,$$

which is nearest, on the right, to $\vartheta_k^{(n)}$.

4.

In this section we consider the number of roots of polynomials in a given interval. We already mentioned in the introduction that if

$$\overline{\lim_{n\to\infty}}\, [|\, l_\nu(x)\, |]^{1/n} \leqq 1 \tag{69}$$

uniformly for $-1 \leqq x \leqq +1$ and $\nu = 1, 2, \cdots, n$, then the fundamental points of $\mathfrak{M}$ are uniformly distributed. We present an elementary proof for this Fejérian-theorem. Here and also later a theorem of M. Riesz[21] plays a most important part, so—because of its shortness—we reproduce it as

Lemma XI. *If a trigonometric polynomial of order n, $f(\vartheta)$, takes its absolute maximum in $[0, 2\pi]$ at $\vartheta = \vartheta_0$, then there is no root of $f(\vartheta)$ in* $\left[\vartheta_0 - \dfrac{\pi}{2n}, \vartheta_0 + \dfrac{\pi}{2n}\right]$.

Suppose the theorem to be untrue. Without loss of generality let $\vartheta_0 = 0$, $f(0) = 1$ and suppose, that the nearest root, the distance of which is less than $\pi/2n$, lies to the right. But in this case the curves $y = f(\vartheta)$ and $y = \cos n\vartheta$ would have at $\vartheta = 0$ at least a double point of intersection and by the premise a third one in $[0, \pi/n]$. In any of the intervals $[\pi/n, 2\pi/n]$, $[2\pi/n, 3\pi/n]$, $\cdots$, $[(2n - 2)\pi/n, (2n - 1)\pi/n]$ they would also have at least one intersection; hence the trigonometric polynomial $f(\vartheta) - \cos n\vartheta$ of order n had in $[0, 2\pi]$ $(2n + 1)$ roots, which is impossible.

Corollary. *If a trigonometric polynomial of order n takes its absolute maximum between two real roots, the distance of these roots cannot be less than π/n. The statement holds for all n.*

Theorem XII. *If upon the matrix* $\mathfrak{M}$

$$[|\, l_k(x)\, |]^{1/n} \leqq 1 + \epsilon, \qquad k = 1, 2, \cdots n, \qquad -1 \leqq x \leqq +1,$$

for $n > n_{10}(\epsilon)$, then we have for any $0 \leqq \alpha < \beta \leqq \pi$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\substack{\nu \\ \alpha \leqq \vartheta_\nu^{(n)} \leqq \beta}} 1 = \frac{\beta - \alpha}{\pi}.$$

[21] M. Riesz: *Eine trigonometrische Interpolationsformel usw.*, Jahresbericht der deutschen Mathematischerver, 1915, pp. 354–368.

PROOF. If the theorem would be untrue, we had in $[0, \pi]$ a subinterval $[\alpha, \beta]$ and a c_{63} such, that there would be an infinity of integers $n_1 < n_2 < \cdots$ for which the number of the n_k^{th} fundamental points lying in $[\alpha, \beta]$ is less than $\frac{1}{\pi}(\beta - \alpha - c_{63})n_k$. We may assume $c_{63} < \frac{\beta - \alpha}{6}$ and write instead of n_k simply n. Let

$$\varphi_k = \alpha + k\frac{\pi}{n+1} \tag{70a}$$

where k runs over the integers (positive, negative and 0), for which φ_k lies in $\left[0, \alpha + \frac{c_{63}}{4}\right]$; for $n \to \infty$ the number of these φ_k asymptotically equals $\frac{1}{\pi}\left(\alpha + \frac{c_{63}}{4}\right)n$. Let further

$$\psi_l = \beta + l\frac{\pi}{n+1}, \tag{70b}$$

where l runs over the integers (positive, negative and 0) for which ψ_l lies in $\left[\beta - \frac{c_{63}}{4}, \pi\right]$; for $n \to \infty$ the number of these asymptotically equals $\frac{1}{\pi}\left(\pi - \beta + \frac{c_{63}}{4}\right)n$. Let further

$$G(x) \equiv \prod_{\nu}{}' (x - \cos \vartheta_\nu^{(n)}) \prod_k (x - \cos \varphi_k) \prod_l (x - \cos \psi_l), \tag{71}$$

where $\prod'$ is to be extended over the $\vartheta_k^{(n)}$ lying in $[\alpha, \beta]$. The degree of $G(x)$ is by the premise and the definition of φ_k and ψ_l, for sufficiently large n,

$$< \frac{1}{\pi}(\beta - \alpha - c_{63})n + \frac{1}{\pi}\left(\alpha + \frac{c_{63}}{3}\right)n + \frac{1}{\pi}\left(\pi - \beta + \frac{c_{63}}{3}\right)n = \left(1 - \frac{c_{63}}{3\pi}\right)n. \tag{72}$$

As the order of the trigonometric polynomial $G(\cos \vartheta)$ is less than n and the distance of its consecutive roots in $\left[0, \alpha + \frac{c_{63}}{4}\right]$ and $\left[\beta - \frac{c_{63}}{4}, \pi\right]$ is less than π/n, $G(\cos \vartheta)$ takes, by lemma XI, its absolute maximum in $\left[\alpha + \frac{c_{63}}{4}, \beta - \frac{c_{63}}{4}\right]$, at a place $\vartheta = \gamma$ say. Let finally

$$F(x) = G(x)\left\{1 - \frac{(x - \cos \gamma)^2}{4}\right\}^{[c_{63}n/8\pi]}, \tag{73}$$

where the brackets in the exponent denote the greatest integer contained. Then, by (72) the degree of $F(x)$ is

$$< \left(1 - \frac{c_{63}}{3\pi}\right)n + \frac{c_{63}}{4\pi}n = \left(1 - \frac{c_{63}}{12\pi}\right)n \tag{74}$$

and, like $G(x)$, $F(x)$ takes its absolute maximum at $\vartheta = \gamma$, too. Then

$$F(x) = \sum_{\nu=1}^{n} F(x_\nu^{(n)}) l_\nu(x)$$

i.e. for $x = \cos\gamma$

$$|G(\cos\gamma)| = |F(\cos\gamma)| = \left|\sum_{\nu=1}^{n} F(x_\nu^{(n)}) l_\nu(\cos\gamma)\right| = \left|\sum_\nu{}'' F(x_\nu^{(n)}) l_\nu(\cos\gamma)\right|,$$

where, by definition of $F(x)$, $\sum''$ refers only to those $\vartheta_\nu^{(n)}$, which are not in $[\alpha, \beta]$. But then, by the hypothesis, for $n > n_{10}(\epsilon)$

$$|G(\cos\gamma)| < (1+\epsilon)^n \sum_\nu{}'' |F(x_\nu^{(n)})|$$

$$< (1+\epsilon)^n \sum_\nu{}'' |G(x_\nu^{(n)})| \left(1 - \frac{(x_\nu^{(n)} - \cos\gamma)^2}{4}\right)^{c_{63}n/8\pi}.$$

As by definition of γ

$$|G(\cos\gamma)| \geqq |G(x_\nu^{(n)})|, \qquad \nu = 1, 2, \cdots n,$$

we have *a fortiori*

$$1 \leqq (1+\epsilon)^n \sum_\nu{}'' \left(1 - \frac{(x_\nu^{(n)} - \cos\gamma)^2}{4}\right)^{c_{63}n/8\pi}$$

$$< n(1+\epsilon)^n \max\left[\left(1 - \frac{\left(\cos\alpha - \cos\left(\alpha + \frac{c_{63}}{4}\right)\right)^2}{4}\right)^{c_{63}n/8\pi}, \left(1 - \frac{\left(\cos\left(\beta - \frac{c_{63}}{4}\right) - \cos\beta\right)^2}{4}\right)^{c_{63}n/8\pi}\right],$$

which is, with sufficiently small ϵ, untrue for $n > n_{11}(\epsilon)$ and thus the theorem is proved.

By theorem XII and lemma VII we obtain

THEOREM XIII. *Let $p(x)$ be in $[-1, +1]$ non-negative, and L-integrable further suppose that its roots form an aggregate of measure 0; then for the roots of the n^{th} polynomial $\cos\vartheta_\nu^{(n)}$ belonging to $p(x)$ we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\substack{\nu \\ \alpha \leqq \vartheta_\nu^{(n)} \leqq \beta}} 1 = \frac{\beta - \alpha}{\pi},$$

where $[\alpha, \beta]$ denotes any fixed subinterval of $[0, \pi]$.

If we want to secure uniform distribution with error-term, we require the following Fejérian

THEOREM XIV. *If for a matrix*

$$|l_\nu(x)| \leqq D, \qquad -1 \leqq x \leqq +1, \qquad \nu = 1, 2, \cdots n, \qquad n = 1, 2, \cdots,$$

then for the elements of the n^{th} *row,* $\cos \vartheta_\nu^{(n)}$ $(\nu = 1, 2, \cdots, n)$ *and for any subinterval* $[\alpha, \beta]$ *of* $[0, \pi]$ *satisfying* $(\beta - \alpha)n \geqq c_{69}(D, \epsilon)$ *we have*

$$\left| \sum_{\alpha \leqq \vartheta_\nu^{(n)} \leqq \beta} 1 - \frac{\beta - \alpha}{\pi} n \right| < c_{70}(D, \epsilon)\{(\beta - \alpha)n\}^{\frac{1}{2}+\epsilon},$$

where we emphasize that $c_{70}(D, \epsilon)$ *is independent of* α *and* β*, too.*

PROOF. Consider first the upper estimate. Let $[\alpha, \beta]$ be a fixed subinterval—without loss of generality we may suppose $\beta - \alpha \leqq \frac{\pi}{4}$ —$\left[\frac{\beta - \alpha}{\pi} n\right] = k$ and assume

$$(75) \qquad \sum_{\alpha \leqq \vartheta_\nu^{(n)} \leqq \beta} 1 = k + l.$$

From α rightwards let us cut off k-times the distance π/n and leftwards $[\frac{1}{4}l]$-times until we reach A, further from β rightwards also $[\frac{1}{4}l]$-times until we reach B; if $\alpha - \pi/n\,[\frac{1}{4}l] \leqq 0$ or $\beta + \pi/n\,[\frac{1}{4}l] \geqq \pi$, set $A = 0$, correspondingly $B = \pi$. Let the points of division be φ_ν. Let further

$$(76) \qquad G_1(x) = \prod_\nu (x - \cos \varphi_\nu) \prod_\mu{}' (x - \cos \vartheta_\mu^{(n)}),$$

where $\prod'$ in the second product runs over the $\vartheta_\nu^{(n)}$ lying outside of $[\alpha, \beta]$. In this case $G_1(\cos \vartheta)$ is a pure cosine-polynomial, whose order $\leqq n - k - l + k + 2[\frac{1}{4}l] < n - \frac{1}{2}l$. Then the distance of two consecutive roots of $G_1(\cos \vartheta)$ in $[A, B]$ is not greater than π/n, i.e., by lemma XI $G_1(\cos \vartheta)$ takes its absolute maximum outside of this interval, at a point $\vartheta = \lambda$, say. Let

$$(77) \qquad F_1(x) = G_1(x) T_{[\frac{1}{4}l]}\left(-1 + 2\,\frac{x - \cos \beta}{\cos \alpha - \cos \beta}\right),$$

where $T_{[\frac{1}{4}l]}(\cos \varphi) = \cos\,[\frac{1}{4}l]\varphi$. As the degree of $F_1(x)$ is not greater, than $n - \frac{1}{4}l$, we represent $F_1(x)$ by the Lagrange-interpolatory polynomial formed upon the n^{th} row of the matrix $\mathfrak{M}$ and obtain

$$F_1(x) \equiv \sum_{\nu=1}^{n} F_1(x_\nu^{(n)}) l_\nu(x)$$

i.e.

$$(78) \qquad |F_1(\cos \lambda)| = \left| \sum_{\nu=1}^{n} F_1(x_\nu^{(n)}) l_\nu(\cos \lambda) \right| = \left| \sum_\nu{}'' F_1(x_\nu^{(n)}) l_\nu(\cos \lambda) \right|,$$

where $\sum''$ runs over integers, for which $\vartheta_\nu^{(n)}$ lies in $[\alpha, \beta]$. As for any $1 \leqq \nu \leqq n$

$$|G_1(\cos \lambda)| \geqq |G_1(x_\nu^{(n)})|,$$

by (78) and the hypothesis we have

$$
(79)\qquad \begin{aligned} &\left| T_{[\frac{1}{2}l]}\left(-1 + 2\,\frac{\cos\lambda - \cos\beta}{\cos\alpha - \cos\beta}\right)\right| \\ &\qquad \leqq D \sum_{\alpha \leqq \vartheta_\nu^{(n)} \leqq \beta} \left| T_{[\frac{1}{2}l]}\left(-1 + 2\,\frac{\cos\vartheta_\nu^{(n)} - \cos\beta}{\cos\alpha - \cos\beta}\right)\right|. \end{aligned}
$$

Each term of the right-hand-side-sum is not greater than 1 and the number of terms is less[8] than $c_{71}(D)k$, i.e. by (79)

$$
(80)\qquad \left| T_{[\frac{1}{2}l]}\left(-1 + 2\,\frac{\cos\lambda - \cos\beta}{\cos\alpha - \cos\beta}\right)\right| \leqq c_{72}(D)\cdot k.
$$

If A or B fall upon one of the borders of $[0, \pi]$ λ cannot be there, thus from λ it may be assumed

$$
\min\left(|\lambda - \alpha|, |\lambda - \beta|\right) \geqq \frac{\pi}{n}\left[\frac{l}{4}\right].
$$

Without loss of generality we may suppose $\lambda \geqq \beta + [\frac{1}{4}l]\,\pi/n$. Then we have

$$
(81)\qquad \left| -1 + 2\,\frac{\cos\lambda - \cos\beta}{\cos\alpha - \cos\beta}\right| > 1 + 2\,\frac{\cos\beta - \cos\left(\beta + \dfrac{\pi}{n}\left[\dfrac{l}{4}\right]\right)}{\cos\alpha - \cos\beta}.
$$

Suppose $0 \leqq \alpha_1 < \alpha_2 < \alpha_3 \leqq \pi$, $\alpha_3 - \alpha_2 = \delta_1$, $\alpha_2 - \alpha_1 = \delta_2$; then we obviously have

$$
(82)\qquad \begin{aligned} \frac{\cos\alpha_2 - \cos\alpha_3}{\cos\alpha_1 - \cos\alpha_2} &= \frac{\sin\dfrac{\delta_1}{2}\sin\left(\alpha_2 + \dfrac{\delta_1}{2}\right)}{\sin\dfrac{\delta_2}{2}\sin\left(\alpha_2 - \dfrac{\delta_2}{2}\right)} \\ &\geqq \frac{\sin^2\dfrac{\delta_1}{2}}{\sin\dfrac{\delta_2}{2}\sin\left(\delta_1 + \dfrac{\delta_2}{2}\right)} > c_{73}\,\frac{\delta_1^2}{\delta_2(\delta_1 + \delta_2)}. \end{aligned}
$$

Case I. $l \geqq k \geqq 20$. We apply (82) with $\alpha_1 = \alpha$, $\alpha_2 = \beta$, $\alpha_3 = \beta + [\frac{1}{4}l]\pi/n$; then we have

$$
\delta_1 \geqq \frac{\pi}{n}\left(\frac{l}{4} - 1\right) \geqq \frac{\pi}{n}\left(\frac{k}{4} - 1\right) \geqq \frac{k\pi}{5n} > \frac{1}{6}\,\delta_2
$$

and

$$
(83)\qquad \frac{\cos\beta - \cos\left(\beta + \left[\dfrac{l}{4}\right]\dfrac{\pi}{n}\right)}{\cos\alpha - \cos\beta} > \frac{c_{73}}{7}\,\frac{\delta_1}{\delta_2} > c_{75}\,\frac{l}{k} > c_{75}.
$$

As $T_\nu(x)$ increases monotonely for $x \geqq 1$ and satisfies for $\rho \geqq 0$ the inequality

$$
(84)\qquad T_\nu(1 + \rho) \geqq \tfrac{1}{2}\left(1 + \sqrt{(2\rho)}\right)^\nu,
$$

we have from (80), (81), (83) and (84)

$$c_{76}^{\frac{1}{2}l} \leqq c_{72}(D)\, k$$

$$l < c_{77}(D) \log k$$

which together with $k \leqq l$ gives $k \leqq k_0(D)$.

CASE II. $l < k,\, k \geqq k_0(D)$. In this case $\delta_1 \leqq \dfrac{l\pi}{4n} < \dfrac{k\pi}{4n} < \dfrac{\delta_2}{4}$ and

$$\frac{\cos\beta - \cos\left(\beta + \dfrac{\pi}{n}\left[\dfrac{l}{4}\right]\right)}{\cos\alpha - \cos\beta} \geqq c_{78}\left(\frac{\delta_1}{\delta_2}\right)^2 > c_{79}\left(\frac{l}{k}\right)^2.$$

From (90), (81) and (84) we have

$$\frac{1}{2}\left(1 + 2\sqrt{c_{79}}\,\frac{l}{k}\right)^{[\frac{1}{4}l]} \leqq T_{[\frac{1}{4}l]}\left(1 + 2c_{79}\left(\frac{l}{k}\right)^2\right) < c_{72}(D)k$$

$$l < c_{80}(D)\ (k \log k)^{\frac{1}{2}}. \quad \text{Q.e.d.}$$

Let us now consider the lower estimate; when again $\left[\dfrac{\beta - \alpha}{\pi}\, n\right] = k$, let

$$\sum_{\alpha \leqq \vartheta_\nu^{(n)} \leqq \beta} 1 = k - l. \tag{85}$$

Thus we have to estimate l from above. Now cut off from α to 0 leftward distances of the length $\dfrac{\pi}{n}\left(1 + \dfrac{1}{10k^{\frac{1}{2}-\epsilon}}\right)$, as many times as possible and rightward $[\frac{1}{4}l]$-times until A'; furthermore, from β rightward to π as many times, as possible and leftwards $[\frac{1}{4}l]$-times until B'. As

$$2\left[\frac{l}{4}\right]\frac{\pi}{n} \leqq \frac{l\pi}{2n} < \frac{k\pi}{2n} \leqq \frac{\beta - \alpha}{2},$$

we have $A' < B'$. We denote the points of division by φ_μ'. Let

$$G_2(x) = \prod_\mu (x - \cos\varphi_\mu') \cdot \prod_\kappa (x - \cos\vartheta_\kappa^{(n)}), \tag{86}$$

where the second product refers to the $\vartheta_\kappa^{(n)}$'s lying inside of $[\alpha, \beta]$. In this case $G_2(\cos\vartheta)$ is a pure cosine-polynomial of order

$$k - l + 2\left[\frac{l}{4}\right] + \left[\frac{n\alpha}{\pi\left(1 + \dfrac{1}{10k^{\frac{1}{2}-\epsilon}}\right)}\right] + \left[\frac{n(\pi - \beta)}{\pi\left(1 + \dfrac{1}{10k^{\frac{1}{2}-\epsilon}}\right)}\right]$$

$$< k - \frac{l}{2} + \frac{n}{\pi\left(1 + \dfrac{1}{10k^{\frac{1}{2}-\epsilon}}\right)}(\pi - (\beta - \alpha))$$

$$< k - \frac{l}{2} + \left(\frac{10nk^{\frac{1}{2}-\epsilon}}{\pi(10k^{\frac{1}{2}-\epsilon} + 1)}\right)\left(\pi - \frac{k\pi}{n}\right)$$

$$= n + \frac{n}{10k^{\frac{1}{2}-\epsilon}} \cdot \frac{10k^{\frac{1}{2}-\epsilon}}{1 + 10k^{\frac{1}{2}-\epsilon}} + \frac{k^{\frac{1}{2}+\epsilon}}{10 + k^{\epsilon-\frac{1}{2}}} - \frac{l}{2} < n - \frac{n}{10k^{\frac{1}{2}-\epsilon} + 1} + \frac{k^{\frac{1}{2}+\epsilon}}{10} - \frac{l}{2}$$

Suppose that $l > \frac{1}{5}k^{\frac{1}{2}+\epsilon}$ is true; then the order of $G_2(\cos\vartheta)$ would be

(87) $$< n - \frac{n}{10k^{\frac{1}{2}-\epsilon} + 1}$$

and after lemma XI the place $\vartheta = \lambda$, on which $G_2(\cos\vartheta)$ takes its absolute maximum could be only in a root interval, whose length is

$$\geqq \frac{\pi}{\text{degree of } G_2(\cos\vartheta)} > \frac{\pi}{n}\left(1 + \frac{1}{10k^{\frac{1}{2}-\epsilon}}\right).$$

But on the outside of $[A', B']$ the distance of the consecutive roots of $G_2(\cos\vartheta)$ is $\leqq \frac{\pi}{n}\left(1 + \frac{1}{10k^{\frac{1}{2}-\epsilon}}\right)$; so λ could be only in $[A', B']$. Without loss of generality let $0 \leqq \lambda \leqq \frac{\pi}{2}$; then we have for $0 \leqq \vartheta \leqq \pi$

$$\left|\frac{1}{\sin\dfrac{\vartheta+\lambda}{2}}\right| < \frac{c_{81}}{|\vartheta-\lambda|}, \qquad \left|\frac{1}{\sin\dfrac{\vartheta-\lambda}{2}}\right| < \frac{c_{81}}{|\vartheta-\lambda|}.$$

With this c_{81} let M be the greatest odd number not exceeding $400\, c_{81} \dfrac{n}{10k^{\frac{1}{2}+\epsilon}+1}$ $+ 1$, $N = 2\left[\dfrac{k^{2\epsilon}}{800c_{81}}\right]$

(88a) $$\psi(\cos\vartheta) = \frac{1}{M^N}\left\{\left(\frac{\sin M\dfrac{\vartheta+\lambda}{2}}{\sin\dfrac{\vartheta+\lambda}{2}}\right)^N + \left(\frac{\sin M\dfrac{\vartheta-\lambda}{2}}{\sin\dfrac{\vartheta-\lambda}{2}}\right)^N\right\}$$

and finally

(88b) $$F_2(x) = G_2(x)\psi(x).$$

The degree of the polynomial $\psi(x)$ is $\dfrac{M-1}{2}N < \dfrac{200c_{81}n}{10k^{\frac{1}{2}+\epsilon}}\dfrac{k^{2\epsilon}}{400c_{81}} = \dfrac{n}{20k^{\frac{1}{2}-\epsilon}} <$ $\dfrac{n}{10k^{\frac{1}{2}-\epsilon}+1}$ and according to (87) and (886) the degree of $F_2(x)$ is $\leqq n-1$, i.e.

(89) $$|F_2(\cos\lambda)| = \left|\sum_{\nu=1}^{n} F_2(x_\nu^{(n)})l_\nu(\cos\lambda)\right| = \left|\sum_\nu{}' F_2(x_\nu^{(n)})l_\nu(\cos\lambda)\right| < D\sum_\nu{}' |F_2(x_\nu^{(n)})|,$$

the last two summations refer to the $\vartheta_\nu^{(n)}$'s lying outside of $[\alpha, \beta]$. As $|G_2(\cos\lambda)| \geqq |G_2(x_\nu^{(n)})|$ $(\nu = 1, 2, \cdots, n)$, we should have from (89)

(90) $$|\psi(\cos\lambda)| < D\sum_\nu{}' |\psi(\cos\vartheta_\nu^{(n)})|.$$

It is easy to see that

$$\psi(\cos\lambda) \geqq 1, \qquad |\psi(\cos\vartheta)| < 2\left(\frac{c_{81}}{M(\vartheta - \lambda)}\right)^N,$$

and comparing this with (90) and with our theorem, we have

$$1 < 2D\left(\frac{c_{81}}{M}\right)^N \sum_\nu{}' \frac{1}{(\vartheta_\nu^{(n)} - \lambda)^N}$$

$$(91) \quad < 2D\left(\frac{c_{81}}{M}\right)^N \left[\frac{1}{(\alpha-\lambda)^N} + \frac{1}{\left(\alpha - \frac{c_{82}\pi}{5n} - \lambda\right)^N} + \frac{1}{\left(\alpha - 2\frac{c_{82}\pi}{5n} - \lambda\right)^N} + \cdots \right.$$
$$\left. + \frac{1}{(\beta-\lambda)^N} + \frac{1}{\left(\beta + \frac{c_{82}\pi}{5n} - \lambda\right)^N} + \frac{1}{\left(\beta + 2\frac{c_{82}\pi}{5n} - \lambda\right)^N} + \cdots \right].$$

Bur for $l > c_{83}$ we have

$$|\lambda - \alpha| > \frac{l\pi}{5n}, \qquad |\lambda - \beta| > \frac{l\pi}{5n}$$

and (91) gives $k > c_{84}$

$$(92) \quad 1 < 4D\left(\frac{5c_{81}n}{\pi M}\right)^N \left[\frac{1}{l^N} + \frac{1}{(l+c_{82})^N} + \frac{1}{(l+2c_{82})^N} + \cdots\right] < \frac{4D}{c_{82}}\left(\frac{10c_{81}n}{\pi M l}\right)^N k$$

further for $l > c_{85}$ we have

$$M > \frac{40c_{81}n}{k^{\frac{1}{2}+\epsilon}}, \qquad N > \frac{k^{2\epsilon}}{800c_{81}},$$

which gives from (92)

$$1 < \frac{4D}{c_{82}} k \left(\frac{k^{\frac{1}{2}+\epsilon}}{4\pi l}\right)^N < \frac{4D}{c_{82}} k \left(\frac{5}{4\pi}\right)^{k^{2\epsilon}/800c_{81}}.$$

This means a contradiction for $l > c_{86}(D, \epsilon)$. Q.c.d.

THEOREM XV. *If upon the matrix* $\mathfrak{M}$

$$|l_k(x)| \leqq c_{87}n^{c_{88}}, \qquad -1 \leqq x \leqq +1, \qquad k = 1, 2, \cdots, n, \qquad n = 1, 2, \cdots,$$

then for any subinterval $[\alpha, \beta]$ *of* $[0, \pi]$

$$\left| \sum_{\substack{\nu \\ \alpha \leqq \vartheta_\nu^{(n)} \leqq \beta}} 1 - \frac{\beta - \alpha}{\pi} n \right| < c_{89}(c_{87}, c_{88}, \epsilon) n^{\frac{1}{2}+\epsilon}.$$

PROOF. It will be sufficient to prove that for any subinterval the upper estimate holds, as in this case the respective application for the intervals $[0, \alpha]$ and $[\beta, \pi]$ leads to

$$\left| \sum_{\substack{\nu \\ 0 \leqq \vartheta_\nu^{(n)} \leqq \alpha}} 1 - \frac{\alpha}{\pi} n \right| < c_{89} n^{\frac{1}{2}+\epsilon}$$

and respectively to

$$\left| \sum_{\beta \leq \vartheta_\nu^{(n)} \leq \pi} 1 - \frac{\pi - \beta}{\pi} n \right| < c_{89} n^{\frac{1}{2}+\epsilon}$$

i.e.

$$\sum_{\alpha \leq \vartheta_\nu^{(n)} \leq \beta} 1 = n - \sum_{0 \leq \vartheta_\nu^{(n)} \leq \alpha} 1 - \sum_{\beta \leq \vartheta_\nu^{(n)} \leq \pi} 1 > \frac{\beta - \alpha}{\pi} n - 2c_{89} n^{\frac{1}{2}+\epsilon},$$

which establishes the lower estimate. The proof of the upper estimate is completely analogous to that applied in theorem XIV.

For a sequence of strongly normal polynomials, theorem XIV immediately presents the uniform distribution of roots in $[0, \pi]$ with the error-terms mentioned, but we do not state this in a separate theorem. For orthogonal polynomials according to (50) and to the first remark appended to theorem VIII we may state that if throughout $[-1, +1]$ we have $p(x) \geqq m > 0$ and L-integrable, then

$$| l_\nu(x) | \leqq \left[\frac{1}{m} \int_{-1}^{1} p(t)\, dt \right]^{\frac{1}{2}} \cdot n, \quad \nu = 1, 2, \cdots n, \quad n = 1, 2, \cdots, \quad -1 \leqq x \leqq +1,$$

or respectively, if in $[-1, +1]$ is $p(x)$ L-integrable and $m \leqq p(x) \sqrt{(1 - x^2)} \leqq M$ then

$$| l_\nu(x) | \leqq \left[c_{66} \frac{M}{m} \right]^{\frac{1}{2}}, \quad -1 \leqq x \leqq +1, \quad \nu = 1, 2, \cdots n, \quad n = 1, 2, \cdots.$$

Hence theorem XV and XIV are applicable and we obtain following two theorems:

THEOREM XVI. *If the weight function is L-integrable and satisfies in* $[-1, +1]$ $p(x) \geqq m > 0$ *and the roots of the* n^{th} *orthogonal polynomial are* $\cos \vartheta_\nu^{(n)}$, *then for any* $[\alpha, \beta]$ *of* $[0, \pi]$ *we have*

$$\left| \sum_{\alpha \leq \vartheta_\nu^{(n)} \leq \beta} 1 - \frac{\beta - \alpha}{\pi} n \right| < c_{90}(p, \epsilon) n^{\frac{1}{2}+\epsilon}.$$

THEOREM XVII. *If the weight function* $p(x)$ *is L-integrable and satisfies in* $[-1, +1]$ $0 < m \leqq p(x) \sqrt{(1 - x^2)} \leqq M$, *then for the roots of the* n^{th} *orthogonal polynomial* $\cos \vartheta_\nu^{(n)}$ *and for any subinterval* $[\alpha, \beta]$ *of* $[0, \pi]$ *we have*

$$\left| \sum_{\alpha \leq \vartheta_\nu^{(n)} \leq \beta} 1 - \frac{\beta - \alpha}{\pi} n \right| < c_{91}(p, \epsilon) \{ (\beta - \alpha) n \}^{\frac{1}{2}+\epsilon}$$

if $n(\beta - \alpha) > c_{92}(p, \epsilon)$.

INSTITUTE FOR ADVANCED STUDY, PRINCETON, N. J.;
BUDAPEST, HUNGARY.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

A REMARK ON A PRECEDING PAPER BY J. VON NEUMANN

BY ISRAEL HALPERIN

(Received August 11, 1939)

In a preceding paper 'On rings of operators, III'[1] J. von Neumann has developed a theory of normed operators for operators in an arbitrary factor $\mathfrak{M}$.[2] In part the procedure is as follows: First $[[A]]$, the norm of A, is defined for all $A \epsilon \mathfrak{M}$ of finite rank (Definition 1.3.1) and fundamental sequences of such operators and an equivalence relation for fundamental sequences are defined (Definitions 1.4.1, 1.4.2 and 1.4.3). Then the key theorem (Theorem III) is proved that every fundamental sequence is strongly convergent operatorially, that two fundamental sequences converge strongly to the same limit if and only if they are equivalent, and finally that the norms of a fundamental sequence converge numerically to a limit depending only on the operator limit of the sequence.

The proof of Theorem III makes use of two important lemmas:

LEMMA 1.4.1. *If A_n, $n = 1, 2, \cdots$ is a fundamental sequence with $[[A_n]] \to 0$ as $n \to \infty$, then A_n converges strongly to 0 as $n \to \infty$.*

LEMMA 1.4.3. *If A_n, $n = 1, 2, \cdots$ is a fundamental sequence with A_n converging strongly to 0 as $n \to \infty$, then $[[A_n]] \to 0$ as $n \to \infty$.*

The purpose of the present note is to point out how the proofs of these two lemmas can be simplified by using more fully the results of the paper 'On rings of operators, II' by Murray and von Neumann.[3] We require the following lemma.

LEMMA. *If $\mathfrak{M}$ is a factor in a finite case and $A_n \epsilon \mathfrak{M}$ with A_n converging strongly to 0 as $n \to \infty$, then $[[A_n]] \to 0$ as $n \to \infty$.*

PROOF. There is a finite set of elements $g_1, \cdots, g_m$ such that $[[A_n]]^2 = \mathrm{Tr}_{\mathfrak{M}}(A_n^* A_n) = \sum_{i=1}^m (A_n^* A_n g_i, g_i) = \sum_{i=1}^m (A_n g_i, A_n g_i) = \sum_{i=1}^m \| A_n g_i \|^2$ for all n.[4] Since $A_n g_i \to 0$ as $n \to \infty$ for all i, and m is finite, the lemma follows.

PROOF OF LEMMA 1.4.3. Suppose the lemma false, if possible. Then for some $\epsilon > 0$ we could assume that $[[A_n]] > \epsilon$ for all n, by choice of a suitable subsequence. For some fixed n_0, $[[A_n - A_r]] < \epsilon/2$ for all $n, r \geqq n_0$. From the definition of a fundamental sequence, A_{n_0} is contained in some closed linear manifold $\mathfrak{M}$ of finite dimensionality relative to $\mathfrak{M}$. Let E be the projection on $\mathfrak{M}$.

[1] These *Annals*, vol. 41 no. 1 (January 1940), pp. 94-161.

[2] The writer is greatly indebted to Professor von Neumann for the opportunity to see the paper in manuscript and for the suggestion to publish this note.

[3] In particular the use of Lemma 1.3.2 of 'On rings of operators, III' will be avoided.

[4] This uses Theorem III of 'On rings of operators, II'.

Then EA_nE, $n = 1, 2, \cdots$ is a fundamental sequence converging strongly to 0 as $n \to \infty$, in the factor $\mathfrak{M}_{(\mathfrak{M})}$ which is in a finite case. The preceding lemma now implies that $[[EA_nE]] \to 0$ as $n \to \infty$. But

$$\begin{aligned}[][[EA_nE]] &\geqq [[EA_{n_0}E]] - [[EA_{n_0}E - EA_nE]] \\ &\geqq [[A_{n_0}]] - |||\, E \,|||\, [[A_{n_0} - A_n]]\, |||\, E \,||| \\ &\geqq \epsilon - \epsilon/2 = \epsilon/2\end{aligned}$$

for all $n \geqq n_0$. This contradiction shows that Lemma 1.4.3 cannot be false.

PROOF OF LEMMA 1.4.1. The factor $\mathfrak{M}$ of operators in the Hilbert space $\mathfrak{H}$ is in one of the classes (I), (II) or (III). If $\mathfrak{M}$ is in class (III) the only fundamental sequence is that with $A_n \equiv 0$ and hence the lemma is trivially true. If $\mathfrak{M}$ is in class (I) or (II) we can consider $\mathfrak{H}$ as a sum of closed linear manifolds $\mathfrak{H} = \sum_r^{\oplus} \mathfrak{M}_r$[5] such that Theorem II of 'On rings of operators, II' applies to every $\mathfrak{M}_{(\mathfrak{M}_r)}$.[6] Thus for every r there is a fixed element $g_r \in \mathfrak{M}_r$ such that

$$\mathrm{Tr}_{\mathfrak{M}}(A) = (Ag_r, g_r) \quad \text{for all} \quad A \in \mathfrak{M}_{(\mathfrak{M}_r)} \quad \text{and} \quad [Bg_r\, ; B \in \mathfrak{M}'] = \mathfrak{M}_r .$$

Now let A_n, $n = 1, 2, \cdots$ be a fundamental sequence with $[[A_n]] \to 0$ as $n \to \infty$. Let E_r be the projection on $\mathfrak{M}_r$. Then

$$\begin{aligned}\| A_n g_r \|^2 &= (A_nE_rg_r, A_nE_rg_r) = (E_rA_n^*A_nE_rg_r, g_r) \\ &= \mathrm{Tr}_{\mathfrak{M}}(E_rA_n^*A_nE_r) \leqq \mathrm{Tr}_{\mathfrak{M}}(A_n^*A_n) = [[A_n]]^2.\end{aligned}$$

It follows that $A_ng_r \to 0$ as $n \to \infty$, for every r. Since $A_n(Bg_r) = B(A_ng_r)$ for all $B \in \mathfrak{M}'$, and such B are bounded, it follows that $A_n(Bg_r) \to 0$ as $n \to \infty$. Since these Bg_r and their finite sums are everywhere dense in $\mathfrak{H}$, and since the A_n are uniformly bounded, it follows that $A_nf \to 0$ for all $f \in \mathfrak{H}$, which proves the lemma.

HARVARD UNIVERSITY.

---

5 The number of summands may be finite or countably infinite.
6 This uses Lemma 11.4.3 of 'On rings of operators'.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

ON REGULARLY CONVEX SETS IN THE SPACE CONJUGATE TO A BANACH SPACE

BY M. KREIN AND V. ŠMULIAN

(Received June 6, 1939)

INTRODUCTION

Here and throughout this paper we denote by E a Banach space (a linear, normed, and complete space) and by E^* its conjugate space.[1]

A set $K \subset E^*$ will be called *regularly convex* if for every $g \notin K$ $(g \in E^*)$ there exists an element $x_0 \in E$ such that

$$\operatorname*{Sup}_{f \in K} f(x_0) < g(x_0).$$

The space E^* itself we consider as regularly convex.

Any closed sphere $\sigma(f_0; \rho)$ $(f_0 \in E^*, \rho > 0)$, i.e. the set of all elements $f \in E^*$ such that $|f - f_0| \leqq \rho$ gives us one of the simplest examples of a regularly closed set. We obtain a still more simple example if we consider the set of all $f \in E^*$ such that $c_1 \leqq f(y) \leqq c_2$ where y is an arbitrary element from E and $c_1 \leqq c_2$ any real constants.

Observe that if K is a linear subspace of E^* then the expression regularly convex denotes the same as regularly closed (see Banach [2], p. 116).

It is easily seen that only convex sets in E^* may be regularly convex. Furthermore, every regularly convex set is always closed and moreover weakly closed in the sense of weak convergence of functionals.

It is also clear that the intersection of regularly convex sets taken in arbitrary number is empty or is also a regularly convex set. In view of this fact, for every set $S \subset E^*$ there exists the smallest regularly convex set containing the set S (namely, the intersection of all regularly convex sets containing S). This set K will be called[2] the *regularly convex envelope* of S.

In the first chapter of our paper we state some general properties of regularly convex sets, in particular, in §1 we study the properties of bounded regularly convex sets. Theorem 1, which gives a representation of elements of the regularly convex envelope of a bounded set by the elements of the set itself, is the central theorem of this paragraph.

On the basis of this theorem we show that the linear aggregate and the convex envelope of a finite number of regularly convex bounded sets are also regularly convex. Most of the theorems of §1 may also be obtained by developing some

[1] The terminology used will be that of Banach's well known book [2]. Numerals in brackets refer to the list of references at the end of the paper.

[2] In analogy with the definitions of linear envelope or convex envelope of a set.

transfinite methods proposed first by Banach (see [2] p. 116–122). Establishing a simple lemma we have succeeded in finding a quite elementary method of proving all these theorems. Formulated in a more particular form, this lemma has been used in the classical moment problem long ago (see M. Riesz [12]; Achyeser-Krein [1], Article 2; I. J. Schoenberg [14]; L. Kantorovitch [5]). But it seems that in this paper the lemma is for the first time applied to the study of regularly convex sets. In §2 we extend a part of our results to regularly convex unbounded sets. The fundamental result of this paragraph is that a set is regularly convex if and only if its intersection with every bounded regularly convex set is regularly convex.

The proof of this theorem is based on an idea of Banach's as in §1 Banach's transfinite methods are not used. Only after having established all the main properties of regularly convex sets do we prove in §3 that a set is regularly convex if and only if it is convex and transfinitely closed. This criterion is a generalization of that of Banach's that gives the conditions under which a linear subspace of E^* is regularly closed (convex). The idea of our proof differs somewhat from that of Banach. Furthermore, we show that it is useful to define transfinitely closedness in a somewhat different way than it was defined by Banach. In §4 it is shown that the criterion of regular convexity of a set may be considerably simplified if the space E or the set itself are separable.

In chapter II we show that the study of factor-spaces introduced by Banach ([2], p. 232) and Hausdorff [4] is closely related to that of regularly convex sets. For instance, it is shown that a linear set belonging to E^* is regularly closed (convex) if and only if it is the conjugate space to a factor-space E/G ($G \subset E$). By means of similar propositions we state that if G is a linear subspace of E then $(G^*)^*$ is a linear subspace of $(E^*)^*$ (see §1); if E is regular ($(E^*)^* = E$) then so are G and E/G, and conversely (see §2). If the unit sphere of E is weakly compact, then the spaces G and E/G both have the same property, and conversely (see §3).

In chapter III we study the regularly convex envelopes of weakly compact sets. As has been remarked by S. Mazur [9], the convex envelope of a compact set of E is compact. Using considerably more complicated means, we state the analogous propositions for sets from E or E^* which are weakly compact in any sense. The case in which the set belongs to E has been already discussed in another place (see M. Krein [7] and V. Smulian [17]). The results concerning this case are now obtained from a more general standpoint. The theorems on convex envelopes of weakly compact sets permit us to formulate the fixed point theorems of Schauder in a somewhat more general form than formulated by Schauder [12].

Chapter I

1. Bounded regularly convex sets. Let S be an arbitrary set. Denote by Q_S the linear space of all real-valued bounded functions $\varphi(s)$ defined over S

with the usual definition of the norm

$$\|\varphi\| = \operatorname*{Sup}_{s \in S} |\varphi(s)|.$$

An additive homogeneous (distributive) functional $\mathfrak{p}$ defined over Q_S shall be said to be *nonnegative*, if $\varphi(s) \geqq 0$ $(s \in S)$ implies $\mathfrak{p}(\varphi) \geqq 0$.[3] Such a functional is always continuous, i.e. belongs to Q_S^* (Q_S^* is the conjugate space to Q_S). In fact, for every $\varphi \in Q_S$ we have

$$\|\varphi\| \cdot \varepsilon(s) \pm \varphi(s) \geqq 0$$

where $\varepsilon(s) \equiv 1$ $(s \in S)$. Therefore, $\mathfrak{p}$ being nonnegative we obtain $\|\varphi\| \cdot \mathfrak{p}(\varepsilon) \pm \mathfrak{p}(\varphi) \geqq 0$, and consequently

$$\|\mathfrak{p}\| = \operatorname{Sup} \frac{|\mathfrak{p}(\varphi)|}{\|\varphi\|} = \mathfrak{p}(\varepsilon) < \infty.$$

A nonnegative functional $\mathfrak{p}$ will be said to be normalized if $\|\mathfrak{p}\| = 1$ $(\mathfrak{p}(\varepsilon) = 1)$. The set of all nonnegative normalized functionals $\mathfrak{p} \in Q_S^*$ will be denoted by $\mathfrak{P}_S$.

Lemma.[4] *Let L be a linear subspace of Q_S and let $\varepsilon(s) \in L$. If a distributive functional $\mathfrak{p}_0$ defined over L takes nonnegative values over nonnegative functions $\varphi \in L$, then there exists a nonnegative functional $\mathfrak{p} \in Q_S^*$ such that $\mathfrak{p}(\varphi) = \mathfrak{p}_0(\varphi)$ for all $\varphi \in L$.*

Proof. Let $\psi_1(s) \in Q_S - L$. Denote by $\varphi'(s) \in Q_S$ (or $\varphi''(s) \in Q_S$) any function such that for all $s \in S$

$$(1) \qquad \varphi'(s) \leqq \psi_1(s) \quad (\text{or } \psi_1(s) \leqq \varphi''(s)).$$

Such functions exist for $\varepsilon(s) \in L$ and $-\|\psi_1\| \cdot \varepsilon(s) \leqq \psi_1(s) \leqq \|\psi_1\| \cdot \varepsilon(s)$.

It follows from (1) that always $\mathfrak{p}(\varphi') \leqq \mathfrak{p}(\varphi'')$ and therefore that there exists such a ξ such that

$$(2) \qquad \operatorname{Sup} \mathfrak{p}(\varphi') \leqq \xi \leqq \operatorname{Inf} \mathfrak{p}(\varphi'').$$

Consider the linear space $L_1 \supset L$ formed by all functions ψ of the form: $\psi(s) = \varphi(s) + t\psi_1(s)$, where $\varphi \in L$ and t is an arbitrary real number. Put $\mathfrak{p}_1(\psi) = \mathfrak{p}(\varphi) + t\xi$. The functional $\mathfrak{p}_1$ defined over L_1, is obviously distributive. It is also easy to see that in virtue of (2), the functional $\mathfrak{p}_1$ takes nonnegative values over nonnegative functions $\psi \in L_1$.

Thus we see that it suffices to well order the set $Q_S - L$ in order to arrive by the above described extension of the functional $\mathfrak{p}_0(\varphi)$ to the functional $\mathfrak{p}(\varphi)$ looked for.

---

[3] It may be easily shown that if an additive functional $\mathfrak{p}$ has the property that $\varphi(s) \geqq 0$ implies $\mathfrak{p}(\varphi) \geqq 0$, then it is necessarily homogeneous and consequently distributive, i.e. $\mathfrak{p}(\lambda\varphi + \mu\psi) = \lambda\mathfrak{p}(\varphi) + \mu\mathfrak{p}(\psi)$ $(\varphi, \psi \in Q_s)$.

[4] This lemma is a particular form of a more general one (see Krein [6]). As was indicated in the introduction, this lemma has its origin in the classical moment problem.

From now on S will denote a bounded set in E^*. In this case to every $x \in E$ corresponds a function $\varphi_x(f) \in Q_S$ defined as follows $\varphi_x(f) = f(x)$ $(f \in S)$. Obviously $\varphi_{ax+by}(f) = a\varphi_x(f) + b\varphi_y(f)$. Now every $\mathfrak{p} \in Q_S^*$ generates a functional $g(x)$ $(x \in E)$ by the formula $g(x) = \mathfrak{p}(\varphi_x)$. Obviously $g \in E^*$ since $g(x + y) = g(x) + g(y)$ and $| g(x) | \leqq || \mathfrak{p} || \cdot \operatorname{Sup}_{f \in S} | f(x) | < C | x |$. We shall write for g, $g = \mathfrak{p}(S)$.

THEOREM 1. *Let S be a bounded set in E^*. Then the regularly convex envelope K of S coincides with the set of all elements $g \in E^*$ which may be obtained by the formula*

$$g = \mathfrak{p}(S) \tag{3}$$

where $\mathfrak{p} \in \mathfrak{P}_S$.

PROOF. Denote by K_1 the set of all $g \in E^*$ which may be obtained by formula (3).

The set K_1 contains the set K. In fact, to every point $f_0 \in S$ corresponds the functional $\mathfrak{p}_{f_0} \in \mathfrak{P}_S$ defined as follows $\mathfrak{p}_{f_0}(\varphi) = \varphi(f_0)$ $(\varphi \in Q_S)$, and for which we obviously have $\mathfrak{p}_{f_0}(S) = f_0$.

The set K_1 is regularly closed. In fact, let $g_0 \notin K_1$. Denote by L the linear subspace consisting of all functions $\psi(f) \in Q_S$ of the form $\psi(f) = \varphi_x(f) + c\,\varepsilon(f)$ $(x \in E, -\infty < c < \infty)$. Put $\mathfrak{p}_0(\psi) = g_0(x) + c$. The functional $\mathfrak{p}_0$ defined over L in general may be many-valued. We state that there exists at least one function

$$\psi_0(f) = f(x_0) + c_0 \geqq 0 \qquad (f \in S) \tag{4}$$

such that

$$\mathfrak{p}_0(\psi_0) = g_0(x_0) + c_0 < 0. \tag{5}$$

In fact, admitting the contrary, $\psi(f) \geqq 0$ $(f \in S)$ implies $\mathfrak{p}_0(\psi) \geqq 0$ and, in particular, if $\psi(f) \equiv 0$ $(f \in S)$ then $\mathfrak{p}_0(\psi) = 0$, which shows that $\mathfrak{p}_0$ is single-valued. As $\mathfrak{p}_0$ is also additive and $\varepsilon(f) \in L$, there exists by the Lemma a non-negative functional $\mathfrak{p} \in Q_S^*$ such that $\mathfrak{p}(\psi) = \mathfrak{p}_0(\psi)$ $(\psi \in L)$. But then $| \mathfrak{p} | = \mathfrak{p}(\varepsilon) = \mathfrak{p}_0(\varepsilon) = 1$, i.e. $\mathfrak{p} \in \mathfrak{P}_S$ and furthermore $g_0(x) = \mathfrak{p}_0(\varphi_x) = \mathfrak{p}(\varphi_x)$ $(x \in E)$, i.e. $g_0 = \mathfrak{p}(S) \in K_1$, which contradicts our supposition.

Thus there exists an element $x_0 \in E$ and a constant c_0 such that (4) and (5) hold and consequently

$$\operatorname{Sup}_{f \in S} f(-x_0) \leqq c_0 < g_0(-x_0),$$

which shows that K_1 is regularly closed.

Thus K_1 being regularly closed and containing S, contains also K. It remains to show that every element of K_1 belongs to K. Admitting the contrary, let $g_0 \in K_1$ and $\notin K$. Then K being regularly closed there exists an element x_0 such that $\operatorname{Sup}_{f \in K} f(x_0) < g_0(x_0)$. On the other hand as $g_0 \in K_1$, there exists a

$\mathfrak{p}_0 \in \mathfrak{P}_S$ such that $g = \mathfrak{p}_0(S)$. Therefore

$$g_0(x_0) = \mathfrak{p}_0(f(x_0)) \leqq \operatorname*{Sup}_{f \in S} f(x_0) \leqq \operatorname*{Sup}_{f \in K} f(x_0) < g_0(x_0),$$

which is impossible. The theorem is thus proved.

COROLLARY 1. *For every* $x \in E$

$$\operatorname*{Sup}_{f \in S} f(x) = \operatorname*{Sup}_{g \in K} g(x).$$

In fact, if $g \in K$ then $g = \mathfrak{p}(S)$ ($\mathfrak{p} \in \mathfrak{P}_S$) and therefore

$$g(x) = \mathfrak{p}(\varphi_x) \leqq \operatorname*{Sup}_{f \in S} \varphi_x(f) = \operatorname*{Sup}_{f \in S} f(x).$$

COROLLARY 2. *If for any pair* $x_1 \in E$, $x_2 \in E$, $f(x_1) = f(x_2)$ *for* $f \in S$, *then* $g(x_1) = g(x_2)$ *for* $g \in K$.

In fact, if $g = \mathfrak{p}(S)$ ($\mathfrak{p} \in \mathfrak{P}_S$) then $g(x_1) = \mathfrak{p}(\varphi_{x_1}) = \mathfrak{p}(\varphi_{x_2}) = g(x_2)$ for $\varphi_{x_1}(f) = f(x_1) = f(x_2) = \varphi_{x_2}(f)$.

REMARK. Let R_S be a linear subspace of Q_S containing the function $\varepsilon(f) \equiv 1$ and all functions $\varphi_x(f)$ ($x \in E$). Denote by $\mathfrak{P}_S^R$ the set of all functionals $\mathfrak{p} \in R_S^*$ which are positive (i.e., take nonnegative values over nonnegative functions $\varphi \in R_S$) and have the unit norm ($\| \mathfrak{p} \| = 1$). It is easy to see that theorem 2 remains true if we replace in it the set $\mathfrak{P}_S$ by the set $\mathfrak{P}_S^R$.

Before passing to the next theorem we will consider an example applying theorem 2 in the form indicated in the Remark.

Let S be a sequence $\{f_i\}_1^\infty \in E^*$ converging weakly to an element $f_0 \in E^*$ ($\lim_{i \to \infty} f_i(x) = f_0(x)$ for every $x \in E$). In this case the space Q_S is equivalent to the space (m), for every function $\varphi(f) \in Q_S$ is determined by its values $\xi_1 = \varphi(f_i)$ ($i = 1, 2, \cdots$) and $\| \varphi \| = \operatorname*{Sup}_{f \in S} | \varphi(f) | = \operatorname*{Sup}_{1 \leq i \leq \infty} | \xi_i |$.

The set R of all functions $\varphi \in Q_S$ such that

$$\lim_{i \to \infty} \varphi(f_i)$$

exists, form a linear subspace of Q_S which evidently is equivalent to the space (c). Therefore to every functional $\mathfrak{p} \in R^*$ (see Banach [2], p. 65–67) corresponds a sequence $\{c_i\}_0^\infty$ such that

$$\mathfrak{p}(\varphi) = \sum_{i=1}^{\infty} c_i \varphi(f_i) + c_0 \lim_{i \to \infty} \varphi(f_i) \qquad (\varphi \in R),$$

hence $\| \mathfrak{p} \| = \sum_{i=0}^{\infty} | c_i |$ and $c_i \geqq 0$ ($i = 1, 2, \cdots$) if the functional $\mathfrak{p}$ is positive.

Since the function $\varepsilon(f)$ and every function $\varphi_x(f) = f(x)$ ($x \in E$) belong to R we can apply to S theorem 2 (in extended form), and so we find that the *regularly convex envelope* K *of the set* $S = \{f_i\}_1^\infty$ *consists of all elements* $g \in E^*$, *which admit the representation*

$$(*) \qquad g(x) = \sum_0^{\infty} c_i f_i(x) \qquad (x \in E),$$

where

$$(**) \qquad c_i \geqq 0 \quad (i = 1, 2, \cdots) \quad \text{and} \quad \sum_0^\infty c_i = 1.$$

As (*) and (**) imply that $g = \sum_0^\infty c_i f_i$ we conclude that the regularly convex envelope K of the set $S = \{f_i\}_1^\infty$ coincides with its convex closed envelope.

THEOREM 2. *Let $S_1, S_2, \cdots, S_n$ be a system of bounded sets in E^* and let $K_1, K_2, \cdots, K_n$ be the corresponding system of the regularly convex envelopes of S_i $(i = 1, 2, \cdots, n, K_i \supset S_i)$. Then the regularly convex envelope K of the logical sum $S = S_1 \cup S_2 \cup \cdots \cup S_n$[5] coincides with the convex envelope K' of the sum $K_1 \cup K_2 \cup \cdots \cup K_n$.*

PROOF. Obviously K being convex and containing all K_i contains also K^1. It remains to show that, conversely, $K \subset K^1$. In proving this we may suppose without loss of generality that all sets S_i $(i = 1, 2, \cdots, n)$ are disjoint, for in the contrary case we replace the sets S_i by sets S_i' defined as follows:

$$S_1' = S_1, \qquad S_k' = S_k - S_kS_1 - S_kS_2 - \cdots - S_kS_{k-1} \quad (k = 2, 3, \cdots, n).$$

Take any $\mathfrak{p} \in \mathfrak{P}_S$. Our theorem will be proved if we show that $g = \mathfrak{p}(S)$ belongs to the convex envelope of the sum $K' = K_1 \cup K_2 \cup \cdots K_n$.

Put $\mu_i = \mathfrak{p}(\varepsilon_i)$ $(i = 1, 2, \cdots, n)$, where

$$(6) \qquad \varepsilon_i(f) = \begin{cases} 1 & \text{for } f \in S_i \\ 0 & \text{for } f \in S - S_i \end{cases} \qquad (i = 1, 2, \cdots, n).$$

Obviously $\mu_i \geqq 0$ $(i = 1, 2, \cdots, n)$ and $\mu_1 + \mu_2 + \cdots + \mu_n = 1$.

To every function $\varphi_i \in Q_{S_i}$ corresponds a function φ determined as follows:

$$\varphi(f) = \begin{cases} \varphi(f) & \text{for } f \in S_i \\ 0 & \text{for } f \in S - S_i. \end{cases}$$

Put $\psi_i(\varphi_i) = \mathfrak{p}(\varphi)$. Thus our functional $\mathfrak{p} \in \mathfrak{P}_S$ generates n functionals $\psi_i \in Q_{S_i}^*$ $(i = 1, 2, \cdots, n)$, which, obviously, are nonnegative and have respectively the norms $\| \psi_i \| = \mathfrak{p}(\varepsilon_i(f)) = \mu_i$ $(i = 1, 2, \cdots, n)$.

Define for each $\mu_i > 0$ the functional $\mathfrak{p}_i \in \mathfrak{P}_S$ by the formula

$$\mathfrak{p}_i = \frac{1}{\mu_i} \psi_i.$$

Observe now that each function $\varphi \in Q_S$ generates a function in each of Q_{S_i} $(i = 1, 2, \ldots, n)$. We shall denote this function by the same letter φ. As for every $\varphi \in Q_S$

$$\varphi(f) = \sum_1^n \varepsilon_i(f)\varphi(f)$$

[5] I.e. the union of all elements belonging to at least one of the sets $S_1, \cdots, S_n$.

we obtain

$$\mathfrak{p}(\varphi) = \sum_1^n \mathfrak{p}(\varepsilon_i \varphi) = \sum_{\mu_i > 0} \mu_i \mathfrak{p}_i(\varphi).$$

In particular, as $g(x) = \mathfrak{p}(f(x))$ $(x \in E)$, we have

$$g(x) = \sum_{\mu_i > 0} \mu_i \mathfrak{p}_i(f(x)) = \sum_{\mu_i > 0} \mu_i g_i(x) \qquad (x \in E),$$

where $g_i = \mathfrak{p}_i(S_i) \in K_i$.

Consequently, recalling that $\sum_{\mu_i > 0} \mu_i = 1$, we conclude that $g = \sum \mu_i g_i$ belongs to the convex envelope of the sum $K_1 \cup K_2 \cup \cdots \cup K_n$, which completes our proof.

COROLLARY. *If $K_1, K_2, \cdots, K_n$ are bounded regularly convex sets in E^*, then their convex envelope is also regularly convex.*

To obtain this corollary we put $S_i = K_i$ $(i = 1, 2, \cdots, n)$ in theorem 2 just proved.

If $a_1, a_2, \cdots, a_n$ are any real numbers we understand by $a_1 S_1 + a_2 S_2 + \cdots + a_n S_n$ the set of all f that admit the representation $f = a_1 f_1 + \cdots + a_n f_n$ where $f_i \in S_i$ $(i = 1, 2, \cdots, n)$.

The method used by proving the last theorem permits also to prove the

THEOREM 3. *Let $S_1, \cdots, S_n$ and $K_1, \cdots, K_n$ have the same meaning as in the preceding theorem. Then the regularly convex envelope $\mathfrak{K}$ of the set $\mathfrak{S} = a_1 S_1 + \cdots + a_n S_n$ coincides with the set $\mathfrak{K} = a_1 K_1 + a_2 K_2 + \cdots + a_n K_n$.*

PROOF. In proving the theorem we may assume all $a_i > 0$ for if some $a_j < 0$ we may replace such an a_j by $-a_j$ and the corresponding K_j by $-K_j$ without changing $\mathfrak{K}$, $\mathfrak{K}'$ and $\mathfrak{S}$; if, furthermore, some $a_l = 0$, then we exclude from our considerations such a_l's and the corresponding K_l.

If we replace every set S_i by the set $S_i + g$ $(i = 1, 2, \cdots, n)$, where $g \in E^*$ is a fixed element, the sets $\mathfrak{K}$ and $\mathfrak{K}'$ are replaced respectively by the sets $\mathfrak{K} + g_1$ and $\mathfrak{K}' + g_1$, where $g_1 = \left(\sum_1^n a_i\right) g$. Therefore we may assume also that every set S_i lies at a positive distance from zero. But then evidently we may choose constants $\lambda_1 > 0, \cdots, \lambda_n > 0$ such that all the sets $S_i' = \lambda_i S_i$ $(i = 1, 2, \cdots, n)$ are disjoint. Hereby

$$\mathfrak{S} = \sum_1^n \frac{a_i}{\lambda_i} S_i', \qquad \mathfrak{K}' = \sum_i^n \frac{a_i}{\lambda_i} K_i',$$

where $K_i' = \lambda_i K_i$ is the regularly convex envelope of the set S_i' $(i = 1, 2, \cdots, n)$. In view of this we may also suppose without loss of generality that the given sets are disjoint.

It is easy to see that $\mathfrak{K} \supset \mathfrak{K}'$. It remains to show that $\mathfrak{K} \subset \mathfrak{K}'$. Put $S = S_1 \cup S_2 \cup \cdots \cup S_n$ and let $\varepsilon_i(f) \in Q_S$, $(i = 1, 2, \cdots, n)$, be the functions defined in (6). Consider now any nonnegative functional $\mathfrak{p} \in Q_S^*$ satisfying the

following conditions

$$\mathfrak{p}(\varepsilon_i) = a_i \qquad (i = 1, 2, \cdots, n). \tag{7}$$

We state that $g = \mathfrak{p}(S) \in \mathfrak{K}'$. In fact, reasoning as above in proving the preceding theorem, we can state that the functional $\mathfrak{p}$ generates n functionals $\mathfrak{p}_i \in \mathfrak{P}_{S_i}$ such that for every $\varphi \in Q_S$

$$\mathfrak{p}(\varphi) = \sum_1^n \mathfrak{p}(\varepsilon_i \varphi) = \sum_1^n a_i \mathfrak{p}_i(\varphi) \tag{8}$$

and therefore putting $g_i = \mathfrak{p}_i(S_i) \in K_i$ $(i = 1, 2, \cdots, n)$, we find

$$g = \sum_1^n a_i g_i \in \mathfrak{K}'.$$

Take now a $g_0 \in E^*$ which does not belong to $\mathfrak{K}'$. Denote by L the linear subspace consisting of all functions $\psi(f) \in Q_S$ of the form $\psi(f) = c_1\varepsilon_1(f) + \cdots + c_n\varepsilon_n(f) + f(x)$, where $x \in E$ and c_i $(i = 1, 2, \cdots, n)$ are arbitrary real numbers. In this space let $\mathfrak{p}_0(\psi) = c_1a_1 + \cdots + c_na_n + g_0(x)$.

The functional $\mathfrak{p}_0$ defined over L may be many-valued. We state that there exists at least one function

$$\psi_0(f) = f(x_0) + c_1^0\varepsilon_1(f) + \cdots + c_n^0\varepsilon_n(f) \geqq 0 \qquad (f \in S) \tag{9}$$

such that

$$\mathfrak{p}_0(\psi_0) = g_0(x_0) + c_1^0a_1 + \cdots + c_n^0a_n < 0. \tag{10}$$

For in the contrary case if a $\psi(f) \geqq 0$ $(f \in S)$ then $\mathfrak{p}_0(\psi) \geqq 0$ and, in particular, $\psi(f) \equiv 0$ and then $\mathfrak{p}_0(\psi) = 0$, that is $\mathfrak{p}_0$ is a single-valued (evidently additive) functional taking over nonnegative functions $\varphi(f) \in L$ nonnegative values. Since furthermore $\varepsilon(f) = \varepsilon_1(f) + \cdots + \varepsilon_n(f)$, we may apply our Lemma and conclude that there exists a nonnegative functional $\mathfrak{p} \in Q_S^*$ such that $\mathfrak{p}(\psi) = \mathfrak{p}_0(\psi)$ for every $\psi \in L$. In particular

$$\mathfrak{p}(\varepsilon_i) = \mathfrak{p}_0(\varepsilon_i) = a_i \qquad (i = 1, 2, \cdots, n) \tag{11}$$

and furthermore $g_0(x) = \mathfrak{p}_0(\varphi_x) = \mathfrak{p}(\varphi_x)$, i.e. $g_0 = \mathfrak{p}(S)$. But by the above $g_0 \in \mathfrak{K}'$, which contradicts our supposition.

Thus there exists such an element $x_0 \in E$ and numbers $c_1^0, \cdots, c_n^0$ such that (9) and (10) hold. Now observe that (9) is equivalent to the system of the following inequalities: $f_i(x_0) + c_i^0 \geqq 0$ for all $f_i \in S_i$ $(i = 1, 2, \cdots, n)$. Multiplying these inequalities respectively by $a_i > 0$ $(i = 1, 2, \cdots, n)$ and adding them up we obtain

$$f(x_0) + a_1c_1^0 + \cdots + a_nc_n^0 \geqq 0 \quad \text{for all} \quad f \in \mathfrak{S}. \tag{12}$$

Therefore comparing (12) and (10) we find

$$\operatorname*{Sup}_{f \in \mathfrak{K}} f(-x_0) = \operatorname*{Sup}_{f \in \mathfrak{S}} f(-x_0) \leqq a_1c_1^0 + \cdots + a_nc_n^0 < g_0(-x_0).$$

Hence $g_0 \notin \mathfrak{K}$. Thus if $g_0 \notin \mathfrak{K}'$, then $g_0 \notin \mathfrak{K}$, i.e. $\mathfrak{K}' \supset \mathfrak{K}$. The theorem is proved.

COROLLARY 1. *If $K_1, K_2, \cdots, K_n$ are bounded regularly convex sets in E^*, then whatever be the numbers $a_1, \cdots, a_n$ the set $a_1K_1 + \cdots + a_nK_n$ is also regularly convex.*

To obtain this corollary we put $S_i = K_i$ $(i = 1, 2, \cdots, n)$ in the theorem just proved.

COROLLARY 2. *Let $K \subset E^*$ be a regularly convex bounded set and $\rho > 0$ an arbitrary fixed number. The logical sum K_ρ of all spheres $\sigma(f, \rho)$ of radius ρ and with center f in K is also regularly convex.*

In fact, $K_\rho = K + \sigma(\theta, \rho)$ and both K and $\sigma(\theta, \rho)$ are regularly closed.

We shall now prove the following theorem.

THEOREM 4. *Let $K_1 \subset E^*$ and $K_2 \subset E^*$ be two disjoint regularly convex bounded sets. Then the distance d between K_1 and K_2 is positive, and to every $d' < d$ $(d' > 0)$ corresponds an $x_0 \in E$ such that*

$$\sup_{f \in K_1} f(x_0) < \inf_{f \in K_2} f(x_0) - d' \,|\, x_0 \,|. \tag{13}$$

PROOF. By corollary I to theorem 3 the set $K = K_1 - K_2$ is regularly convex. As K_1 and K_2 are disjoint, so the element zero, $\theta \subset E$, does not belong to K and consequently lies at a positive distance from K. Obviously, d is also the distance between K_1 and K_2.

By corollary 2 to theorem 3 the set $K_{d'}$ is regularly convex and does not contain the zero θ. Consequently by the definition of regularly convex sets there exists an $x_0 \in E$ such that

$$\sup_{f \in K_{d'}} f(x_0) < 0.$$

Take an $f_0 \in K$, then $\sigma(f_0, d') \subset K_{d'}$; hence

$$\sup_{f \in \sigma(f_0, d')} f(x_0) = f_0(x_0) + d' \,|\, x_0 \,| < 0.$$

Consequently

$$\sup_{f \in K} f(x_0) < -d' \,|\, x_0 \,|. \tag{14}$$

Obviously (14) is equivalent to (13).

2. Unbounded regularly convex sets. We now pass to the study of unbounded regularly convex sets in E^*. First of all we shall state the following theorem, fundamental in the sequel.

THEOREM 5. *In order that an unbounded set $K \subset E^*$ be regularly convex it is necessary and sufficient that the intersection of K with every bounded regularly convex set be also regularly convex.*

PROOF.[6] The necessity of the indicated condition is trivial. It remains to show its sufficiency, that is if K satisfies this condition, then to any given $g_0 \notin K$

[6] The main idea of this proof is borrowed from Banach ([2], pp. 120–121).

corresponds an $x_0 \in E$ such that

$$\operatorname*{Sup}_{f \in K} f(x_0) < g(x_0).$$

Denote by d the distance of g from K. Obviously $d > 0$. Take any sequence $\varepsilon_1 = 1 > \varepsilon_2 > \varepsilon_3 >, \cdots, \varepsilon_n \to 0$ and a $d' < d$ ($d' > 0$). We will now form by recurrence a sequence $\{x_i\}_1^\infty$ possessing the property: if for some $f \in E^*$ and some natural number n the following conditions

$$(15) \qquad f \in K, \qquad f(x_i) \geqq g(x_i) - \frac{d' \,|\, x_i \,|}{\varepsilon_i} \qquad (i = 1, 2, \cdots, n)$$

hold, then

$$(16) \qquad |f - g| > \frac{d'}{\varepsilon_{n+1}}.$$

To obtain the first element x_1 of this sequence we consider the set H_1 of all f such that

$$f \in K \quad \text{and} \quad |f - g| \leqq \frac{d'}{\varepsilon_2}$$

the set H_1 being the intersection of K with the sphere $\sigma(g, d'/\varepsilon_2)$ is by the assumed condition a regularly convex bounded set. As g lies at a distance $\geqq d$ from H_1, there exists such an x_1 that

$$\operatorname*{Sup}_{f \in H_1} f(x_1) < g(x_1) - d' \,|\, x_1 \,|.$$

Obviously, by this choice of x_1 the conditions $f \in K$, $f(x_1) \geqq g(x_1) - d' \,|\, x_1 \,|$ imply $|\, g - f \,| > d'/\varepsilon_2$. Let now $x_1, x_2, \cdots, x_n$ be formed so that (15) implies (16). Then x_{n+1} can be obtained as follows: Consider the set H_{n+1} of all $f \in E^*$ which satisfy the conditions (15) and belong to the sphere

$$(17) \qquad |\, g - f \,| \leqq \frac{d'}{\varepsilon_{n+2}}.$$

If the set H_{n+1} is empty, take for x_{n+1} an arbitrary element different from zero. In another case H_{n+1} being the intersection of K and of $n + 1$ regularly convex sets of which one is bounded (namely, the sphere $\sigma(g, d'/\varepsilon_{n+2})$) is itself regularly convex. As the conditions (15) imply (16), the set H_{n+1} lies at a distance $> d'/\varepsilon_{n+1}$ from g, and therefore there exists an x_{n+1} such that

$$\operatorname*{Sup}_{f \in H_{n+1}} f(x_{n+1}) < g(x_{n+1}) - \frac{d' \,|\, x_{n+1} \,|}{\varepsilon_{n+1}}.$$

Consequently, if

$$(18) \qquad f \in K, \qquad f(x_i) \geqq g(x_i) - \frac{d' \,|\, x_i \,|}{\varepsilon_i} \qquad (i = 1, 2, \cdots, n + 1),$$

then $f \notin H_{n+1}$, which is possible if (17) does not hold, that is $|\, g - f \,| > d'/\varepsilon_{n+2}$.

Thus the existence of the required sequence $\{x_i\}_1^\infty$ is proved. It follows from the definition of this sequence that if $f \in K$ then at least one of the inequalities $f(x_i) \leqq g(x_i) - d' \mid x_i \mid / \varepsilon_i$ $(i = 1, 2, \cdots)$ is true, i.e.

$$\text{(19)} \qquad \operatorname*{Sup}_{1 \leqq i \leqq \infty} \varepsilon_i \frac{g(x_i) - f(x_i)}{\mid x_i \mid} > d'.$$

Putting

$$y_i = \varepsilon_i \frac{x_i}{\mid x_i \mid} \qquad (i = 1, 2, \ldots),$$

we may write the relation (19) in the form:

$$\text{(20)} \qquad \operatorname{Sup} \{g(y_i) - f(y_i)\} > d'.$$

Denote by K_0 the set of all sequences $\{f(y_i)\}_1^\infty$, where $f \in K$. As $\mid y_i \mid = \varepsilon_i \to 0$, we have $\lim_{i \to \infty} f(y_i) = 0$ and therefore $K_0 \subset (c_0)$, where (c_0) is the space of all sequences $\xi = \{\xi_i\}_1^\infty$ converging toward zero with the definition of the norm $\| \xi \| = \operatorname*{Sup}_{1 \leqq i \leqq \infty} \mid \xi_i \mid$.

The inequalities (20) show that the sequence $\{g(y_i)\}_1^\infty \in (c_0)$ is at a distance $\geqq d'$ from K_0. Furthermore, as K is convex, so is K_0.

Therefore there exists in (c_0) a hyperplane

$$\text{(21)} \qquad a_1\xi_1 + a_1\xi_2 + \cdots = a_0 \left(\sum_1^\infty \mid a_i \mid = 1 \right),$$

which passes through the point $\{g_0(y_i)\}_1^\infty$, i.e.

$$\text{(22)} \qquad \sum_1^\infty a_i g_0(y_i) = a_0,$$

and whose distance from K_0 is $\geqq d'$. The last fact can be expressed by a suitable choice of the sign of a_0 as follows:

$$\text{(23)} \qquad \sum_1^\infty a_i f_0(y_i) \leqq a_0 - d' \qquad (f \in K).$$

Putting then

$$\text{(24)} \qquad x_0 = \sum_1^\infty a_i y_i$$

we obtain from (22) and (23) that

$$\text{(25)} \qquad f(x_0) \leqq g(x_0) - d' \qquad (f \in K),$$

which completes the proof of our theorem.

THEOREM 6. *In order that a linear subspace $F \subset E^*$ be regularly closed it is necessary and sufficient that there exists in F a bounded regularly convex set K which has in F interior points (i.e. is in F a convex body).*

PROOF. To show the necessity of the condition, take in F any bounded set S containing in F interior points. Form then the regularly convex envelope K of S. As the intersection of K and F (if F is regularly closed) must be also regularly convex, it coincides with K. Therefore $K \subset F$ satisfies all the requirements of our theorem.

Conversely, let F contain a bounded regularly closed body K. Obviously, we may suppose that the null element $\theta \in K$, for in the contrary case we bring about this situation by a suitable translation of K in F. Let now K' be any bounded regularly closed set. It is clear that for a sufficiently large λ the intersection of K' and F coincides with the intersection D of K' and λK. But K' and λK being regularly closed, so is D, which proves our theorem.

THEOREM 7. *If $K_1, K_2, \cdots, K_n$ are regularly convex sets in E^* and if furthermore all these sets except maybe one are bounded, then the set $K = a_1K_1 + \cdots + a_nK_n$ is regularly convex whatever the constants $a_1, a_2, \cdots, a_n$ maybe.*

PROOF. In the case in which all sets $K_1, \cdots, K_n$ are bounded the theorem coincides with corollary I of theorem 3. Let then $K_1, \cdots, K_{n-1}$ be bounded and K_n unbounded ($a_n \neq 0$). Put

$$C_i = \operatorname*{Sup}_{f \in K_i} |f| \qquad (i = 1, 2, \cdots, n-1).$$

Take now any bounded set $H \subset E^*$:

$$C = \operatorname*{Sup}_{f \in H} |f| < \infty.$$

Choose a number $R > 0$ such that

$$|a_n| R - \sum_1^{n-1} |a_i| C_i > C.$$

Denote by K_n' the intersection of K_n with the sphere $\sigma(\theta, R)$ and put $K' = a_1K_1 + \cdots + a_{n-1}K_{n-1} + a_nK_n'$. The sets $K_1, \cdots, K_{n-1}, K_n'$ being regularly convex and bounded, so is the set K' and, consequently, the intersection $K' \cap H$. But it is easy to see that by our choice of R we have $K \cap H = K' \cap H$. Thus $K \cap H$ is regularly convex, which proves our theorem.

COROLLARY. *Corollary 2 of theorem 3 remains true if K is an unbounded regularly convex set.*

Consequently, theorem 4, being based on the two corollaries of theorem 3, also admits a generalization, namely:

THEOREM 8. *Theorem 4 remains true if one of the sets K_1, K_2 is bounded and another is unbounded.*

Observe also that for the case in which K_1 is unbounded and K_2 consists of only one element g, this theorem is a consequence of the inequality (25), for it follows from (24) that

$$|x_a| \leqq \sum_1^{\infty} |a_i| |y_i| \leqq \sum_1^{\infty} |a_i| = 1$$

and consequently (25) implies

$$\operatorname*{Sup}_{f \epsilon K} f(x_0) \leqq g(x_0) - d' \mid x_0 \mid.$$

3. Transfinite methods. In this paragraph we will outline another method for studying the regularly convex set which is based on some ideas of Banach (see [2] pp. 118–126).

Let ϑ be a limit ordinal. Denote by (ρ_ϑ) the linear space of bounded transfinite sequences of real numbers $X = \{\xi_\alpha\}_1^\vartheta$ of the type ϑ $(1 \leqq \alpha < \vartheta)$, in which the operations of addition and multiplication by a scalar are defined as follows:

1°. *If* $X = \{\xi_\alpha\}$, $Y = \{\eta_\alpha\}$, *then* $X + Y = \{\xi_\alpha + \eta_\alpha\}$

2°. *If* $X = \{\xi_\alpha\}$ *and* λ *is a scalar, then* $\lambda X = \{\lambda \xi_\alpha\}$.

Put

$$p(X) = \overline{\lim_{\alpha \to \vartheta}}\, \xi_\alpha.$$

Obviously, the functional $p(X)$ has the properties:

$$p(X + Y) \leqq p(X) + p(Y), \qquad p(\lambda X) = \lambda p(X), \qquad \lambda \geqq 0.$$

By the well-known theorem of Banach (see [2] p. 29) there exists an additive functional $\mathfrak{p}(X)$, such that

a) $$\mathfrak{p}(X) \leqq p(X) \qquad (X \epsilon (\rho_\vartheta)),$$

and consequently

b) $$-p(-X) \leqq \mathfrak{p}(X) \qquad (X \epsilon (\rho_\vartheta)).$$

Put for arbitrary $X = \{\xi_\alpha\}_1^\vartheta \epsilon (\rho_\vartheta)$

$$\operatorname*{Lim}_{\alpha \to \vartheta} \xi_\alpha = \mathfrak{p}(X).$$

It follows from a) and b) that

$$\underline{\lim_{\alpha \to \vartheta}}\, \xi_\alpha \leqq \operatorname*{Lim}_{\alpha \to \vartheta} \xi_\alpha \leqq \overline{\lim_{\alpha \to \vartheta}}\, \xi_\alpha \qquad (\{\xi_\alpha\}_1^\vartheta \epsilon (P_\vartheta)). \tag{26}$$

Furthermore, as $\mathfrak{p}(X)$ is an additive functional we have

$$\operatorname*{Lim}_{\alpha \to \vartheta} \{\lambda \xi_\alpha + \mu \eta_\alpha\} = \lambda \operatorname*{Lim}_{\alpha \to \vartheta} \xi_\alpha + \mu \operatorname*{Lim}_{\alpha \to \vartheta} \eta_\alpha. \tag{27}$$

From now on in this paragraph we shall agree to associate with every limit ordinal ϑ a fixed operation $\operatorname*{Lim}_{\alpha \to \vartheta}$ on (ρ_ϑ) to the real axis, which possesses the properties (26) and (27).

Let now $\{f_\alpha\}_1^\vartheta$ be any ordered sequence of the type ϑ $(1 \leqq \alpha < \vartheta)$, the elements of which belong to E^* and are bounded in their totality, i.e.

$$\operatorname{Sup}_{1 \leqq \alpha < \vartheta} | f_\alpha | < \infty .$$

As for every $x \epsilon E$ the sequence $\{f_\alpha(x)\}$ is bounded, we may make correspond to the sequence $\{f_\alpha\}_1^\vartheta$ a functional defined as follows:

$$f_0(x) = \operatorname{Lim}_{\alpha \to \xi} f_\alpha(x) \qquad (x \epsilon E).$$

In virtue of (26) and (27) the functional $f_0(x)$ is additive and

$$| f_0(x) | \leqq \operatorname{Sup}_{1 \leqq \alpha < \vartheta} | f_\alpha(x) | \leqq \operatorname{Sup}_{1 \leqq \alpha < \vartheta} | f_\alpha | \cdot | x |.$$

Therefore $f_0 \epsilon E^*$ and

$$| f_0 | \leqq \operatorname{Sup}_{1 \leqq \alpha < \vartheta} | f_\alpha |. \tag{28}$$

Modifying slightly Banach's definition ([2], p. 119), we shall call the functional f_0 *transfinite limit* of the sequence $\{f_\alpha\}$ and we shall write

$$f_0 = \lim_{\alpha \to \vartheta} f_\alpha . \tag{29}$$

Now it is natural to say that a set $K \subset E^*$ is *transfinitely closed* if it contains the transfinite limits of all ordered bounded sequences $\{f_\alpha\}$ belonging to it.

THEOREM 9. *In order that a set $K \subset E^*$ be regularly convex it is necessary and sufficient that it be transfinitely closed.*

PROOF. The condition is necessary. In fact, if a bounded ordered sequence $\{f_\alpha\}$ $(1 \leqq \alpha < \vartheta)$ belongs to the regularly convex set K, then the functional

$$f_0 = \operatorname{Lim}_{\alpha \to \vartheta} f_\alpha$$

also belongs to K for in the contrary case there exists an $x_0 \epsilon E^*$ such that

$$\operatorname{Sup}_{1 \leqq \alpha < \vartheta} f_\alpha(x_0) \leqq \operatorname{Sup}_{f \epsilon K} f(x_0) < f_0(x_0),$$

which is incompatible with (26).

In proving the sufficiency of our condition we may suppose without loss of generality that K is bounded. In fact, if K is transfinitely closed then by (28) so is the intersection of $K^{(n)}$ with the sphere $\sigma(\theta, n)$. On the other hand if we show that $K^{(n)}$ is regularly closed, then so will K be, by theorem 5.

Thus suppose K bounded and transfinitely closed. Denote by K' the regularly convex envelope of K. To prove our theorem it remains to show that $K' = K$. Take a $g \epsilon K'$. Let us well order the elements of E ($E = \{x_\alpha\}$, $(\alpha < \vartheta)$). To prove that $g \epsilon K$ we show that to every $\vartheta_1 \leqq \vartheta$ corresponds at least one $f \epsilon K$ such that

$$f(x_\alpha) = g(x_\alpha) \qquad (\alpha < \vartheta_1). \tag{30}$$

Let first $\vartheta_1 = n + 1$ be an integer. Consider the set $\mathfrak{R}_n \subset R_n$ (R_n is n-dimensional Euclidean space) of all points $(\xi_1, \cdots, \xi_n)$, where

$$\xi_1 = f(x_1), \cdots, \xi_n = f(x_n) \qquad (f \in K). \tag{31}$$

It is easily seen that $\mathfrak{R}_n$ is convex and closed (the latter in virtue of the transfinite closedness of K). We state that the point

$$\eta_1 = g(x_1), \cdots, \eta_n = g(x_n) \tag{32}$$

belongs to $\mathfrak{R}_n$. In fact, in the contrary case there exists by a theorem of Minkowski a system of numbers $\alpha_1, \cdots, \alpha_n$ such that

$$\operatorname*{Sup}_{(\xi_1, \cdots, \xi_n) \in \mathfrak{R}_n} \sum_1^n \alpha_i \xi_i < \sum_1^n \alpha_i \eta_i,$$

which by (31) and (32) may be written as follows

$$\operatorname*{Sup}_{f \in K} f\left(\sum_1^n \alpha_i x_i\right) < g\left(\sum_1^n \alpha_i x_i\right).$$

Since $g \in K'$ this inequality contradicts corollary 1 of theorem 1. Thus $(\eta_1, \cdots, \eta_n) \in \mathfrak{R}_n$, i.e. for $\vartheta_1 = n + 1$ there exists a $f \in K$ such that (30) holds.

Now we can apply the transfinite induction. Indeed let the limit ordinal ϑ_1 have the property: to every $\beta < \vartheta_1$ corresponds a $f_\beta \in K$, such that $f_\beta(x_\alpha) = g(x_\alpha)$ $(\alpha < \beta)$. Then K being transfinitely closed the functional

$$f_{\vartheta_1} = \operatorname*{Lim}_{\beta \to \vartheta_1} f_\beta$$

belongs to K and, furthermore, as it satisfies the conditions $f_{\vartheta_1}(x_\alpha) = g(x_\alpha)$ $(\alpha < \vartheta_1)$, the possibility of applying the transfinite induction is established. This completes the proof of our theorem.

The reader who is familiar with the book of Banach has probably remarked that theorem 9 presents a generalization of Banach's theorem concerning transfinitely closed subspaces ([2], Lemmas 2, 3, pp. 119–122). Our method of proving theorem 9 is slightly different from that of Banach, though it is possible also to prove this theorem following his methods.

It is easy to see that theorem 7 may be obtained as an almost immediate corollary of theorem 9. However, we should meet some difficulties in this deduction of the theorem if we had defined transfinitely closed sets exactly as Banach defined them.

We also find it interesting that all the main properties of regularly convex sets may be obtained as has been shown in the preceding paragraphs without resorting to transfinite limits.

4. Separable sets. The criterion that a set $K \subset E^*$ is regularly convex is much simplified in the cases in which the space E or the set K itself are separable. We begin with

THEOREM 10. *Let E be a separable Banach space. Then a set $K \subset E^*$ is regularly convex if and only if it is convex and weakly closed (as a set of linear functionals).*[7]

PROOF. The necessity of the indicated condition being trivial, we shall prove its sufficiency. Thus let $K \in E^*$ be a convex and weakly closed set. As the sphere $\sigma_n = \sigma(\theta, n)$ is also convex and weakly closed, so is the intersection, K_n, of K and σ_n. Now if we prove that K_n is regularly convex, then so will be the set K (by theorem 5). Therefore in proving that K is regularly convex, we may suppose at once that K is bounded.

Let $g \in K'$, where K' is the regularly convex envelope of the set K. Our theorem will be proved if we show that $g \in K$. Let $\{x_n\}$ be a dense sequence in E. Since E is separable, K is weakly compact. Therefore we may reason the same way as in proving theorem 9 and show that to every $n = 1, 2, \cdots$ corresponds a $f_n \in K$ such that

$$f_n(x_i) = g(x_i) \qquad (i = 1, 2, \cdots, n). \tag{33}$$

K being weakly compact, there exists a subsequence $\{f_{n_\nu}\}$ and an $f_0 \in K$ such that $f_0(x) = \lim f_{n_\nu}(x)$ $(x \in E)$. Then in virtue of (33), $f_0(x_i) = g(x_i)$ $(i = 1, 2, \cdots)$ and consequently $g = f_0 \in K$. Thus the theorem is proved.

One can easily obtain theorem 10 from theorem 9 as Banach had done by proving his particular case (see footnote 7), but we wished to show how the theorem may be proved without transfinite limits.

LEMMA 1. *Whatever be the separable subspace $F \subset E^*$, there exists a separable subspace $G \subset E$ such that $F \subset G^*$.*[8]

PROOF. Let $\{f_m\}_1^\infty$ be a sequence dense in F. To every f_n corresponds a sequence $\{x_{mn}\}_{n=1}^\infty$ such that

$$\lim_{n\to\infty} \frac{|f_m(x_{mn})|}{|x_{mn}|} = |f_m| \qquad (n = 1, 2, \ldots).$$

Denote by G the linear envelope of the set $\{x_{mn}\}$ $(m, n = 1, 2, \cdots)$. It is easy to see that $|f|_E = |f|_G$ if $f \in F$. Therefore we may write $F \subset G^*$.

LEMMA 2. *Let S be a separable bounded set in E^*. If for each sequence $\{f_n\} \subset S$ the functional*

$$f_0 = \operatorname*{Lim}_{n\to\infty} f_n$$

belongs to H, then the set H is weakly compact (in the sense of weak convergence of functionals).[9]

Here we denote by $\operatorname*{Lim}_{n\to\infty}$ an arbitrary fixed operation (generalized limit)

---

[7] For the particular case when K is a linear subspace of E^* one can find this theorem in the book of Banach ([2], p. 124).

[8] See V. Šmulian [15] and [16].

[9] This lemma is taken from a paper of V. Šmulian [16], where it is proved in a somewhat more general form.

corresponding to the first infinite ordinal number $\vartheta = \omega$ which was defined in the preceding paragraph.

PROOF. Take an arbitrary sequence $\{f_n\}_1^\infty \subset S$. Denote by F the linear envelope of the set S. As the set F is separable, there exists by the preceding Lemma a separable subspace $G \subset E$, such that $F \subset G^*$. Let $\{x_i\}_1^\infty$ be a sequence dense in G. It is clear that if $f', f'' \in F$ and $f'(x_i) = f''(x_i)$ $(i = 1, 2, \cdots)$, then $f' = f''$. Choose by the diagonal process of Cantor a subsequence $\{f_{n_\nu}\}_{\nu=1}^\infty$ such that there exists $\lim_{\nu\to\infty} f_{n_\nu}(x_i)$ for each $i = 1, 2, \cdots$.

We now assert that $\lim_{\nu\to\infty} f_{n_\nu}(x)$ exists for every $x \in E$. Admit the contrary, i.e. that for some $x_0 \in E$ the $\lim_{\nu\to\infty} f_{n_\nu}(x_0)$ does not exist. Then we can form two subsequences $\{f_{p_\nu}\}_{\nu=1}^\infty$ and $\{f_{q_\nu}\}_{\nu=1}^\infty$ such that there exist two different $\lim_{\nu\to\infty} f_{p_\nu}(x_0)$ and $\lim_{\nu\to\omega} f_{q_\nu}(x_0)$. On the other hand putting

$$f' = \operatorname*{Lim}_{\nu\to\infty} f_{p_\nu}, \qquad f'' = \operatorname*{Lim}_{\nu\to\omega} f_{q_\nu}$$

we obtain that

$$f'(x_i) = \lim_{\nu\to\infty} f_{p_\nu}(x_i) = \lim_{\nu\to\infty} f_{q_\nu}(x_i) = f''(x_i) \qquad (i = 1, 2, \cdots) \tag{34}$$

and

$$f'(x_0) \neq f''(x_0). \tag{35}$$

Since by the condition of our Lemma $f' \in S \subset F$, $f'' \in S \subset F$, the relations (34) and (35) are in contradiction.

We shall say that a set $S \subset E^*$ is a *locally weakly compact* set if every bounded sequence $\{f_n\} \subset S$ contains a subsequence $\{f_{n_\nu}\}$ which converges weakly to a functional belonging to S.

THEOREM 11. *In order that a separable set K be regularly convex it is necessary and sufficient that K be convex and locally weakly compact.*[10]

PROOF. To prove the necessity of our condition, consider the intersection K_n of K with the sphere $\sigma(\theta, n)$ $(n = 1, 2, \cdots)$. If K is regularly convex then so is K_n. Therefore, by theorem 9, K_n satisfies the conditions of lemma 2 and is consequently weakly compact. Hence K is locally weakly compact.

To prove the sufficiency of our condition we take an arbitrary $g \notin K$. Consider the linear envelope F of the set $\{K, g\}$. By lemma 1, there exists a separable subspace $G \subset E$ such that $F \subset G^*$ and consequently $K \subset G^*$ and $g \in G^*$. K being locally weakly compact as a set of E^*, it is the same regarded as a set of G^*. Consequently K is weakly closed as a set of G^* and, being convex, is by theorem 10 regularly convex as a set of G^*. Hence there exists an element $x_0 \in G$ such that

$$\operatorname*{Sup}_{f \in H} f(x_0) < g(x_0),$$

which completes our proof.

[10] The particular case of this theorem, when K is a linear subspace was already established by V. Šmulian [15].

Chapter II

1. Principal properties of factor-spaces. If G is a linear subspace of E, we denote by E/G the factor group, obtained from E and G if we consider E and G as a group and a subgroup, respectively, with respect to the addition operation of the elements. The elements of E/G will be denoted by Greek letters ξ, η, ζ, $\cdots$. So every element $\xi \epsilon E/G$ can be considered as a set of elements $x \epsilon E$ such that, if $x_0 \epsilon \xi$, then the element x belongs to ξ if and only if $x - x_0 \epsilon G$.

It is easy to see (see Hausdorff [4]) that E/G becomes a Banach space if the multiplication of the elements $\xi \epsilon E/G$ by a scalar and the norm $|\xi|$ are defined as follows:

The product $\lambda\xi$ is the element $\eta \epsilon E/G$ which contains all elements λ where $x \epsilon \xi$; the norm

$$|\xi| = \operatorname*{Inf}_{x \epsilon \xi} |x| = \operatorname*{Inf}_{u \epsilon G} |x_0 - u| \qquad (x_0 \epsilon \xi).$$

If F is a linear subspace of E^*, we may consider also the factor space E^*/F. The elements of the latter we shall denote by Greek letters φ, ψ, χ, $\cdots$.

Lemma. *If G is a closed linear subspace of E, and F*[11] *is the set of all elements $f \epsilon E^*$ such that*

$$f(x) = 0 \qquad (x \epsilon G) \tag{36}$$

then G is the set of all elements $x \epsilon E$ such that,

$$f(x) = 0 \qquad (f \epsilon F). \tag{37}$$

Conversely, if F is a linear, regularly closed subspace of E^*, and G[12] *is the set of all elements $x \epsilon E$ for which* (37) *holds, then F is the set of all elements $f \epsilon E^*$ such that* (36) *holds.*

Proof. To prove the first statement we have to show that if $x_0 \notin G$ then there exists an $f_0 \epsilon F$ such that $f_0(x_0) \neq 0$. But if $x_0 \notin G$, G being closed, there exists an $f_0 \epsilon E^*$ such that $f_0(x_0) = 1$ and $f_0(x) = 0$ for $x \epsilon G$. As the last condition signifies that $f \epsilon F$, the statement is proved. The second part of the lemma follows from the definition of regularly closed subspace of E^*.

Theorem 12. *If F is a regularly closed subspace of E^*, and G denotes the set of all elements $x \epsilon E$ such that $f(x) = 0$ for $f \epsilon F$, then*

$$F = (E/G)^* \quad \text{and} \quad G^* = E^*/F. \tag{38}$$

Proof. To prove the first of these equalities let us put for every $f \epsilon F$, $f(\xi) = f(x)$, if $\xi = \xi_x$ contains x. Obviously, $f(\xi)$ is an additive single-valued functional defined on E/G. As $|f(\xi_x)| = |f(x)| = |f(x - u)| \leqq |f|_E \cdot |x - u|$ for $u \epsilon G$, $|f(\xi_x)| \leqq |f|_E \cdot \operatorname{Inf} |x - u| = |f|_E \cdot |\xi_x|$. Thus $f(\xi)$ is a linear functional defined on $H = E/G$ with the norm

$$|f|_H \leqq |f|_E . \tag{39}$$

[11] Obviously, F is regularly closed.
[12] Obviously, G is closed.

Conversely, let $f(\xi)$ denote a linear functional defined on E/G. Putting $f(x) = f(\xi)$ if $x \in \xi$ we obtain an additive functional defined on E and satisfying the condition

$$f(x) = 0 \qquad (x \in G). \tag{40}$$

As $| f(x) | = | f(\xi) | \leqq | f |_H | \xi | \leqq | f |_H | x | \; (x \in \xi), f \in E^*$, and

$$| f |_E \leqq | f |_H . \tag{41}$$

Further, by (40) and the Lemma $f \in F$. Also every functional $f \in (E/G)^*$ generates a functional denoted by the same letter $f \in F$ and conversely.

Comparing (39) and (41), we see that $| f |_E = | f |_H$, which completes the proof of the first equality (38).

Now let $f_0(x)$ be a linear functional defined on $G (f_0 \in G^*)$. Let φ_0 denote the set of all $f \in E^*$ such that $f(x) = f_0(x)$ for $x \in G$. According to the theorem of Hahn-Banach the set φ_0 is not empty and moreover contains at least one element f' such that $| f' |_E = | f |_G$. If $g_0 \in \varphi_0$, then $g \in \varphi_0 \; (g \in E^*)$ if and only if $g(x) = g_0(x)$ for $x \in G$, i.e. $g - g_0 \in F$. Hence we can consider the set φ_0 as an element of E^*/F. Obviously

$$| \varphi_0 | = \operatorname*{Inf}_{f \in F} | f |_E \leqq | f' |_E = | f_0 |_G .$$

Thus every element $f \in G^*$ generates an element $\varphi = \varphi_f \in E^*/F$ such that

$$| \varphi_f | \leqq | f |_G . \tag{42}$$

Conversely, given an element $\varphi_0 \in E^*/F$, for each pair $g_1 \in \varphi_0$ and $g_2 \in \varphi_0$, we have $g_1(x) = g_2(x)$ for $x \in G$. Consequently, φ_0 determines an element $f_0 \in G^*$ such that $f_0(x) = g(x)$ for $x \in G$ and every $g \in \varphi_0$. Since $| f_0 |_G \leqq | g |_E$ for $g \in \varphi_0$,

$$| f_0 |_G \leqq \operatorname*{Inf}_{g \in F} | g |_E = | \varphi_0 |. \tag{43}$$

Thus every element $\varphi \in E^*/F$ generates an element $f \in G^*$ such that $\varphi = \varphi_f$, and (43) holds. Comparing (42) and (43) we see that $| f |_G = | \varphi_f |$, which completes the proof of the second part of the theorem.

In view of the Lemma we can give another form to our theorem.

THEOREM 12′. *If G is a linear closed subspace of E and F is the set of all functionals $f \in E^*$ such that $f(x) = 0$ for $x \in G$, then $G^* = E/F$ and $F = (E/G)^*$.*

E^*

Using the notations of the last theorem, let us now consider the spaces E^{**} and G^{**} conjugate to the spaces E^* and G^*, respectively. Let L be the set of all elements $\mathfrak{p} \in E^{**}$ such that $\mathfrak{p}(f) = 0$ for $f \in F$. Then according to the last theorem L is the space conjugate to the space E^*/F, which in turn is the conjugate space of G. Thus $G^{**} = L \subset E^{**}$. We obtain thus

THEOREM 13. *If G is a linear subspace of E, then G^{**} may be regarded as a subspace of E^{**}.*

2. 1. The Banach space E is said to be regular if to every $\mathfrak{p} \in E^{**}$ corresponds an $x \in E$ such that $\mathfrak{p}(f) = f(x)$ for $f \in E^*$. A. J. Plessner has remarked that

every linear closed subspace of a regular space is also regular.[13] This proposition is an immediate consequence of the proof of theorem 13. The following theorem complements the proposition of A. J. Plessner.

THEOREM 14. *Let G be a linear closed subspace of E. In order that E be regular it is necessary and sufficient that G and E/G be regular.*

PROOF. Let E be regular. Then, as we already know, G is regular. We have to show that the factor space E/G is also regular. Keeping the notations of the preceding paragraph we may state that the space conjugate to E/G is a subspace of the regular space E^* and consequently is regular. But if $F = (E/G)^*$ is regular, then so is E/G.[14]

Conversely, let G and E/G be regular. Let $\mathfrak{p}_0 \in E^{**}$. Considering the functional $\mathfrak{p}_0(f)$ on $F = (E/G)^*$, we find in view of the regularity of E/G that there exists a $\xi_{x_0} \in E/G$, such that $\mathfrak{p}_0(f) = f(\xi_{x_0}) = f(x_0)$ for $f \in F$.

On the other hand putting $\mathfrak{p}_1(f) = \mathfrak{p}(f) - f(x_0)$ for $f \in E^*$, we obtain that $\mathfrak{p}_1(f) = 0$ for $f \in F$. Therefore, considering $\mathfrak{p}_1$ as a linear functional defined on $E^*/F = G^*$, we find that $\mathfrak{p} \in G^{**}$. Since G is regular, there exists a $x_1 \in G$ such that $\mathfrak{p}_1(\varphi_f) = \varphi_f(x_1)$ for $\varphi_f \in E^*/F$ or $\mathfrak{p}_1(f) = f(x_1)$ for $f \in E^*$. Thus $\mathfrak{p}_0(f) = f(x_0 + x_1) = f(x)$ for $f \in E^*$, which completes the proof of the theorem.

2. A set $\mathfrak{M}$ of a Banach-space E will be said to be *weakly compact* if every sequence $\{x_n\}_1^\infty \subset \mathfrak{M}$ contains a subsequence $\{x_{n_\nu}\}_{\nu=1}^\infty$ which converges *to an element of E.*

In proving the next theorem we shall rely upon two propositions:[15]

A) The unit sphere of a regular Banach space is weakly compact.

B) The unit sphere of a Banach space E is weakly compact if and only if the unit sphere of the conjugate space is weakly compact.

THEOREM 15. *Let G be a linear, closed subspace of E. In order that the unit sphere of the space E be weakly compact it is necessary and sufficient that the unit spheres of the spaces G and E/G be weakly compact.*

PROOF. If the unit sphere of G is weakly compact then it is easy to see that the unit sphere of G is weakly compact. Furthermore, given a bounded sequence $\{\xi_n\}_1^\infty \subset E/G$ ($|\xi_n| < \rho$, $n = 1, 2, \cdots$), we may choose such $x_n \in \xi_n$ so that also $|x_n| < \rho$ ($n = 1, 2, \cdots$). But then there exists a subsequence $\{x_{n_\nu}\}_{\nu=1}^\infty$ that converges weakly to an $x_0 \in E$ and this implies that the sequence $\{\xi_{n_\nu}\}_{\nu=1}^\infty$ converges weakly to ξ_{x_0}.

Conversely, let the unit spheres of G and E/G be weakly compact. In proving that the unit sphere of E is weakly compact it is sufficient to prove the statement in the special case when the space E/G is separable. In fact, given a

[13] This fact is an immediate consequence of the criterion for the regularity of the Banach space which was also indicated by Plessner. A simple proof of this criterion was proposed by V. Gantmacher and V. Šmulian [3].

[14] This follows from the proposition of Plessner: A Banach space is regular if and only if its conjugate space is regular ([8], p. 115).

[15] These propositions have been proved by V. Gantmacher and V. Šmulian [3]. The proposition A), as well as the proposition of Plessner indicated above, have been recently repeated by Pettis [11].

bounded sequence $\{x_n\}_1^\infty \subset E$, we may form the linear, closed envelope E_1 of the set $\{G, x_1, x_2, \cdots\}$. Then E_1/G, is separable and being a subspace of E/G has a weakly compact unit sphere. Therefore if our theorem is established for the case when the factor-space is separable, the unit sphere of E_1 will be weakly compact and consequently the given sequence, $\{x_n\} \subset E_1$, will contain a subsequence $\{x_{n_\nu}\}_{\nu=1}^\infty$ which will converge weakly to an $x_0 \in E_1 \subset F$.

Thus we suppose without loss of generality that the factor-space E/G is separable and consequently regular. On the other hand, to prove that the unit sphere of E is weakly compact it is sufficiently by B) to show that the unit sphere of E^* is weakly compact. To this purpose let us consider the space $F = (E/G)^* \subset E^*$ and $E^*/F = G^*$. The space F is regular, and by A) the unit spheres of F and E^*/F are weakly compact.

But as above by proving that E^* possesses the required property we may suppose that E^*/F is separable and consequently regular. Then, according to theorem 14, E^* is regular and consequently its unit sphere is by A) weakly compact, which completes the proof of the theorem.

3. Linear envelopes of regularly convex sets. The results of this paragraph complete in some respect those of the preceding paragraph.

Let F be the linear envelope of a bounded, regularly convex set $K \subset E^*$. Denote by G the set of all elements x such that $f(x) = 0$ for $f \in K$, and by $(E/G)_K$ the factor group E/G with the new definition of the norm, namely

$$|\xi|_K = \operatorname*{Sup}_{f \in K} |f(x)| \qquad (\xi \in E/G), \tag{44}$$

where x is an arbitrary element belonging to ξ. The set K being bounded, put

$$C = \operatorname*{Sup}_{f \in K} |f|.$$

Then from (44)

$$|\xi|_K \leqq C\,|x|. \tag{45}$$

Moreover since for every $u \in G$

$$|\xi|_K = \operatorname*{Sup}_{f \in K} |f(x - u)| \leqq C|x - u|,$$

we find also that

$$|\xi|_K \leqq C \operatorname*{Inf}_{u \in G} |x - u| = C|\xi|.$$

For every $f \in F: f(x) = f(y)$, if only $x - y \in G$; therefore we can define $f(\xi)$ as equal to $f(x)$, where $x \in \xi$. Obviously, $f(\xi)$ is an additive, single-valued functional defined over $(E/G)_K$. Moreover $f(\xi)$ is continuous over $(E/G)_K$. In fact, every $f \in F$ admits the representation $f(\xi) = \text{const.}\ (f_1(\xi) - f_2(\xi))$, where $f_i \in K$ $(i = 1, 2)$. Since, in virtue of (44), $|f_i(\xi)| \leqq |\xi|_K$ $(i = 1, 2)$, we see that

(46) $$|f|_K = \operatorname{Sup} \frac{|f(\xi)|}{|\xi|_K} < \infty.$$

Conversely, let $f_0(\xi)$ denote a linear functional defined over $(E/G)_K$. Putting $f_0(x) = f_0(\xi)$ if $x \in \xi$ we obtain an additive functional defined over E for which by (45)

(47) $$|f_0(x)| = |f_0(\xi)| \leqq |f_0|_K |\xi|_K \leqq C |f_0|_K |x|.$$

Thus f_0 is a linear functional belonging to E^*. We will show that moreover $f_0 \in F$. To this purpose we denote by K^+ the convex envelope of the sum of the set K and the symmetric set $-K$. By corollary 1 of theorem 2, K^+ is regularly closed. Obviously,

(48) $$|\xi|_K = \operatorname*{Sup}_{f \in K^+} f(x) \qquad (x \in \xi).$$

In proving that $f_0 \in F$ it is sufficient to show that each $f_1 = \lambda f_0 \in K^+$ if $|\lambda| \leqq 1/|f_0|_K$. Admitting for some f_1 the contrary, there will exist an $x_0 \in E$ such that

(49) $$\operatorname*{Sup}_{f \in K^+} f(x_0) < f_1(x_0) = f_1(\xi_0) \leqq |\xi_0|_K \qquad (x_0 \in \xi_0).$$

But this contradicts (48). Thus $f_1 \in K^+$. Observe also that, conversely, if $f \in K$, then $|f|_K = 1$, for according to (44) $|f(\xi)| \leqq |\xi|_K$. Denote by F_K the vector space F in which is introduced the norm $|f|_K$ defined by (46). Then we may formulate the result obtained as follows:

THEOREM 16. *The Banach space F_K is the conjugate space to the space $(E/G)_K$, furthermore, the set K^+ is its unit sphere.*

The space F_K being a conjugate space is complete. On the other hand it follows from (45) that $|f| \leqq C|f|_K$ $(f \in F)$. Therefore in virtue of a well-known theorem of Banach (see [2], p. 41) if F is closed the two norms $|f|$ and $|f|_K$ define the same topology in F, which implies the existence of a constant C_1 such that $|f|_K \leqq C_1 |f|$ $(f \in F)$. Consequently in this case the intersection R of the sphere $\sigma(\theta, 1) \subset E^*$ with F coincides with the intersection of the sphere $\sigma(\theta, 1)$ with the set C_1K. The last two sets being regularly convex, so is the set R. Recalling theorem 6 we arrive at

THEOREM 17. *If the linear envelope F of a bounded, regularly convex set $K \subset E^*$ is closed, then it is regularly closed.*

Note also the following

THEOREM 18. *If $K \subset E^*$ is a bounded regularly convex, separable set, then the factor-space $(E/G)_K$ is separable.*

PROOF. Evidently K^+ is also separable and being bounded and regularly closed is by theorem 11 weakly compact as a set of functionals. Then by theorem 22, which we shall prove in the sequel, the set K^+ is a weakly continuous image of a compact metric set. But K^+ being the unit sphere of the conjugate space to the space $(E/G)_K$, the aforesaid property of K^+ implies (by

bounded sequence $\{x_n\}_1^\infty \subset E$, we may form the linear, closed envelope E_1 of the set $\{G, x_1, x_2, \cdots\}$. Then E_1/G, is separable and being a subspace of E/G has a weakly compact unit sphere. Therefore if our theorem is established for the case when the factor-space is separable, the unit sphere of E_1 will be weakly compact and consequently the given sequence, $\{x_n\} \subset E_1$, will contain a subsequence $\{x_{n_\nu}\}_{\nu=1}^\infty$ which will converge weakly to an $x_0 \epsilon E_1 \subset F$.

Thus we suppose without loss of generality that the factor-space E/G is separable and consequently regular. On the other hand, to prove that the unit sphere of E is weakly compact it is sufficiently by B) to show that the unit sphere of E^* is weakly compact. To this purpose let us consider the space $F = (E/G)^* \subset E^*$ and $E^*/F = G^*$. The space F is regular, and by A) the unit spheres of F and E^*/F are weakly compact.

But as above by proving that E^* possesses the required property we may suppose that E^*/F is separable and consequently regular. Then, according to theorem 14, E^* is regular and consequently its unit sphere is by A) weakly compact, which completes the proof of the theorem.

3. Linear envelopes of regularly convex sets. The results of this paragraph complete in some respect those of the preceding paragraph.

Let F be the linear envelope of a bounded, regularly convex set $K \subset E^*$. Denote by G the set of all elements x such that $f(x) = 0$ for $f \epsilon K$, and by $(E/G)_K$ the factor group E/G with the new definition of the norm, namely

$$|\xi|_K = \operatorname*{Sup}_{f \epsilon K} |f(x)| \qquad (\xi \epsilon E/G), \tag{44}$$

where x is an arbitrary element belonging to ξ. The set K being bounded, put

$$C = \operatorname*{Sup}_{f \epsilon K} |f|.$$

Then from (44)

$$|\xi|_K \leqq C\,|x|. \tag{45}$$

Moreover since for every $u \epsilon G$

$$|\xi|_K = \operatorname*{Sup}_{f \epsilon K} |f(x-u)| \leqq C|x-u|,$$

we find also that

$$|\xi|_K \leqq C \operatorname*{Inf}_{u \epsilon G} |x-u| = C|\xi|.$$

For every $f \epsilon F$: $f(x) = f(y)$, if only $x - y \epsilon G$; therefore we can define $f(\xi)$ as equal to $f(x)$, where $x \epsilon \xi$. Obviously, $f(\xi)$ is an additive, single-valued functional defined over $(E/G)_K$. Moreover $f(\xi)$ is continuous over $(E/G)_K$. In fact, every $f \epsilon F$ admits the representation $f(\xi) = \text{const.}\,(f_1(\xi) - f_2(\xi))$, where $f_i \epsilon K$ $(i = 1, 2)$. Since, in virtue of (44), $|f_i(\xi)| \leqq |\xi|_K$ $(i = 1, 2)$, we see that

$$|f|_K = \operatorname{Sup} \frac{|f(\xi)|}{|\xi|_K} < \infty. \tag{46}$$

Conversely, let $f_0(\xi)$ denote a linear functional defined over $(E/G)_K$. Putting $f_0(x) = f_0(\xi)$ if $x \in \xi$ we obtain an additive functional defined over E for which by (45)

$$|f_0(x)| = |f_0(\xi)| \leqq |f_0|_K |\xi|_K \leqq C |f_0|_K |x|. \tag{47}$$

Thus f_0 is a linear functional belonging to E^*. We will show that moreover $f_0 \in F$. To this purpose we denote by K^+ the convex envelope of the sum of the set K and the symmetric set $-K$. By corollary 1 of theorem 2, K^+ is regularly closed. Obviously,

$$|\xi|_K = \operatorname*{Sup}_{f \in K^+} f(x) \qquad (x \in \xi). \tag{48}$$

In proving that $f_0 \in F$ it is sufficient to show that each $f_1 = \lambda f_0 \in K^+$ if $|\lambda| \leqq 1/|f_0|_K$. Admitting for some f_1 the contrary, there will exist an $x_0 \in E$ such that

$$\operatorname*{Sup}_{f \in K^+} f(x_0) < f_1(x_0) = f_1(\xi_0) \leqq |\xi_0|_K \qquad (x_0 \in \xi_0). \tag{49}$$

But this contradicts (48). Thus $f_1 \in K^+$. Observe also that, conversely, if $f \in K$, then $|f|_K = 1$, for according to (44) $|f(\xi)| \leqq |\xi|_K$. Denote by F_K the vector space F in which is introduced the norm $|f|_K$ defined by (46). Then we may formulate the result obtained as follows:

THEOREM 16. *The Banach space F_K is the conjugate space to the space $(E/G)_K$, furthermore, the set K^+ is its unit sphere.*

The space F_K being a conjugate space is complete. On the other hand it follows from (45) that $|f| \leqq C|f|_K$ $(f \in F)$. Therefore in virtue of a well-known theorem of Banach (see [2], p. 41) if F is closed the two norms $|f|$ and $|f|_K$ define the same topology in F, which implies the existence of a constant C_1 such that $|f|_K \leqq C_1|f|$ $(f \in F)$. Consequently in this case the intersection R of the sphere $\sigma(\theta, 1) \subset E^*$ with F coincides with the intersection of the sphere $\sigma(\theta, 1)$ with the set C_1K. The last two sets being regularly convex, so is the set R. Recalling theorem 6 we arrive at

THEOREM 17. *If the linear envelope F of a bounded, regularly convex set $K \subset E^*$ is closed, then it is regularly closed.*

Note also the following

THEOREM 18. *If $K \subset E^*$ is a bounded regularly convex, separable set, then the factor-space $(E/G)_K$ is separable.*

PROOF. Evidently K^+ is also separable and being bounded and regularly closed is by theorem 11 weakly compact as a set of functionals. Then by theorem 22, which we shall prove in the sequel, the set K^+ is a weakly continuous image of a compact metric set. But K^+ being the unit sphere of the conjugate space to the space $(E/G)_K$, the aforesaid property of K^+ implies (by

the Lemma which will be proved at the beginning of the following paragraph) the separability of the space $(E/G)_K$.

Chapter III

1. Weakly compact sets in the conjugate space. In this paragraph we use the notions *weakly convergent, weakly compact,* etc. for sequences, sets, etc. of elements from E^* in the sense which is usually adopted for functionals. To avoid a misunderstanding we will say also that a set $S \subset E^*$ is *weakly compact in* E^* if every sequence $\{f_n\}_1^\infty \subset S$ contains a subsequence $\{f_{n_\nu}\}$ converging weakly to an element $f_0 \in E^*$; if furthermore, whatever be the sequence $\{f_n\}$, the corresponding elements f_0 all belong to S, then S is simply called *weakly compact* (or *weakly compact in itself*).

We first prove some Lemmas:

Lemma 1. *In order that a Banach space E be separable it is necessary and sufficient that the unit sphere $\sigma = \sigma(\theta, 1)$ of the conjugate space E^* be a weakly continuous image of a compact metric set.*

Proof. If E is separable, then there exists a sequence $\{x_n\}_1^\infty$ dense in the unit sphere $\| x \| \leqq 1$. We introduce a new distance $\rho(g, f)$ between two elements $g \in E^*$ and $f \in E^*$ as follows:

$$\rho(g, f) = \operatorname*{Sup}_{1 \leqq n < \infty} \frac{|f(x_n) - g(x_n)|}{n}. \tag{50}$$

As E is separable, σ is weakly compact. It is easy to see that σ becomes by this new metric a compact metric set, in which the metric convergence coincides with the originally weak convergence. Thus σ can be looked at as a weakly continuous one-to-one image of a compact metric set.

Let now conversely σ be a weakly continuous image of a compact metric set T. Then σ is the range of a weakly continuous function f_t defined over T $(t \in T)$. To every $x \in E$ we make correspond a continuous function $\varphi_x(t)$ $(t \in T)$ as follows: $\varphi_x(t) = f_t(x)$ $(t \in T)$. Thus the space E is isometrically transformed into a linear subspace of the space $[T]$ consisting of all continuous functions $\varphi(t)$ with the usual definition of the norm: $\| \varphi \| = \operatorname*{Sup}_{t \in T} | \varphi(t) |$. But the space $[T]$ is separable, and, consequently, so is E.

Given a set $S \subset E^*$ we denote by $\tilde{S}$ the weak closure of S, i.e. the totality of all elements of S and of all their weak limits.

E^*

Lemma 2. *If E is separable and $S \subset E$ is bounded, then the weak closure $\tilde{S}$ of S is weakly compact and consequently weakly closed.*

Proof. Take in E^* a sphere $\sigma(\theta, r)$, which contains S. In proving the preceding Lemma we have shown, E being separable, that σ can be regarded as a metric compact set in which the metric convergence coincides with the weak convergence. But then $\tilde{S}$ is the metric closure of S in σ and, consequently, is metrically closed and also weakly closed.

From theorem 10 and Lemma 2 follows

THEOREM 19. *If E is separable and $S \subset E^*$ is bounded, then the weak closure $\tilde{S}_{\text{conv.}}$ of the convex envelope $S_{\text{conv.}}$ of S coincides with the regularly convex envelope K of S.*

PROOF. Since K is convex and regularly closed (and consequently weakly closed) we have $\tilde{S}_{\text{conv.}} \subset K$. On the other hand it is evident that $\tilde{S}_{\text{conv.}}$ is convex and, being by lemma 2 weakly closed, it is by theorem 10 regularly convex. Hence $K \subset \tilde{S}_{\text{conv.}}$.

It is now easy to prove the

THEOREM 20. *Let $S \subset E^*$ be a weakly continuous image of a compact metric set T. The weak closure $\tilde{S}_{\text{conv.}}$ of the convex envelope $S_{\text{conv.}}$ of S coincides with the regularly convex envelope K of S. Furthermore, the set $\tilde{S}_{\text{conv.}}$ itself is a weakly continuous one-to-one image of a compact metric set.*

PROOF. Obviously S is a bounded set and therefore K is bounded. Starting from K, introduce in the same way as in Chap. II, §3 the linear subspace $G \subset E$, the factor-space $(E/G)_K$, and the conjugate space $F_K = (E/G)^*_K$. The set S can be regarded as a set in E^* or as a set in F_K. But it is easy to see that in both cases the set $\tilde{S}_{\text{conv.}}$ remains the same. Therefore, recalling theorem 19, we see that our theorem will be proved if we show that the space $(E/G)_K$ is separable. To this end we reason in like manner as in proving Lemma 1. By the condition of our theorem there exists a weakly continuous function f_t, $(t \in T)$, the range of which coincides with S. Make correspond to every $\xi \in (E/G)_K$ the function $\varphi_\xi(t)$ defined as follows: $\varphi_\xi(t) = f_t(x)$ $(t \in T)$, where $x \in \xi$. As in virtue of corollary I of theorem I

$$(51)\qquad |\xi|_K = \operatorname*{Sup}_{x \in K} |f(x)| = \operatorname*{Sup}_{x \in S} |f(x)| = \operatorname*{Sup}_{t \in T} |f_t(x)| = \operatorname*{Sup}_{t \in T} |\varphi_\xi(t)|,$$

the correspondence $\xi \leftrightarrow \varphi_\xi$ between $(E/G)_K$ and the subspace of the space $[T]$ consisting of all continuous functions $\varphi(t)$ defined over T is a linear, isometric one. But the space $[T]$ being separable, so is the space $(E/G)_K$, which completes the proof of our theorem.

THEOREM 21. *Let $S \subset E^*$ be weakly compact in E^*. If $\tilde{S}$ is separable, then it is a weakly continuous one-to-one image of a compact metric set.*

PROOF. Let F be the linear envelope of $\tilde{S}$. Evidently, F is separable and consequently, by Lemma I of Ch. I, §4, there exists a separable subspace $G \subset E$ such that $F \subset G^*$. On the other hand, if G is separable, we can apply Theorem 19 to the set $\tilde{S}$. Thus we obtain that $\tilde{S}$, considered as a subset of G^*, is a weakly continuous one-to-one image of a compact metric set T.

Denote by f_t the image of $t \in T$ in $\tilde{S}$; then $t_n \to t_0$ implies

$$(52)\qquad f_{t_n}(x) \to f_{t_0}(x)$$

for every $x \in G$. Our theorem will be proved if we show that (52) holds for every $x \in E$. Admit the contrary, i.e. that there exists an $x_0 \in E$ such that

$$(53)\qquad \limsup_{n\to\infty} |f_{t_n}(x_0) - f_{t_0}(x_0)| > 0.$$

As S is weakly dense in $\tilde{S}$, the antecedent of S in T is dense in T. Therefore there exist points $t'_n \in T$ $(n = 1, 2, \cdots)$ such that

$$\rho(t_n, t'_n) < \frac{1}{n}, \qquad f_{t'_n} \in S \qquad (n = 1, 2, \ldots).$$

On the other hand, S being weakly compact there exists a subsequence $\{f_{t'_{n_\nu}}\}$ converging weakly to an $f_0 \in \tilde{S}$ as a set of functionals of E^*, that is $\lim_{\nu \to \infty} f_{t'_{n_\nu}}(x) = f_0(x)$ $(x \in E)$. As for every $x \in G$ the function $f_t(x)$ is continuous on the compact T and, consequently, uniformly continuous on T, we have also that

$$\lim_{\nu \to \infty} f_{t_{n_\nu}}(x) = f_0(x) \qquad (x \in G). \tag{54}$$

From (52) and (54) we conclude that $f_0(x) = f_{t_0}(x)$ for every $x \in G$ and, consequently,

$$f_0(x) = f_{t_0}(x) \qquad (x \in E) \tag{55}$$

for $|f' - f''|_E = |f' - f''|_G$, if $f', f'' \in F$. Comparing (53), (54) and (55) we come to a contradiction. Thus the theorem is proved.

Theorem 22. *Let $S \subset E^*$ be weakly compact in E^*. If $\tilde{S}$ is separable, then $\tilde{S}_{\text{conv.}}$ coincides with the regularly convex envelope K of S. Furthermore, $\tilde{S}_{\text{conv.}}$ is a weakly continuous one-to-one image of a compact metric set.*

Proof. As K is weakly closed the set $S' = \tilde{S} \subset K$. In virtue of theorem 21 we can apply to S' theorem 20, thus we obtain that $K = \tilde{S}'_{\text{conv.}}$ is a weakly continuous image of a compact metric set. Since furthermore the set $\tilde{S}'_{\text{conv.}}$ coincides in this case with the set $\tilde{S}_{\text{conv.}}$ the theorem is proved.

2. Weakly compact sets in E. Every element $x \in E$ generates an element $X \in E^{**}$ (E^{**} denotes the conjugate space to the space E^*) by the formula

$$X(f) = f(x) \qquad (f \in E^*).$$

This element X will be denoted by x^{**}.

If a sequence $\{x_n\}_1^\infty$ converges weakly to an element x_0, then the corresponding sequence $\{x_n^{**}\}_1^\infty$ converges weakly (as a set of functionals defined over E^*) to the element x_0^{**}. This permits us to deduce from the theorems which have been proved in the preceding paragraph some properties of weakly compact sets in the space E. The weak compactness of sets in E is understood here in the same sense as in §2, i.e. a set $G \subset E$ will be said to be *weakly compact* in E if every sequence $\{x_n\}_1^\infty \subset G$ contains a subsequence $\{x_{n_\nu}\}$ which converges weakly to an element $x_0 \in E$.

Thus according to this definition if a set $G \subset E$ is weakly compact in E the corresponding set $G^{**} \subset E^{**}$ is weakly compact in E^{**} as a set of functionals defined over E^*, the converse in general not being true.

THEOREM 23. *If $G \subset E$ is a weakly compact set in E its weak closure $\bar{G}$ (i.e. the totality of all elements of G and their weak limits) is weakly compact and weakly closed.*

PROOF. Take an arbitrary sequence $\{x_m\}_1^\infty \subset \bar{G}$. Choose in G for every x_m a sequence $\{x_{mn}\}_{n=1}^\infty$, which converges weakly to x_n ($n = 1, 2, \cdots$). Denote by E_1 the linear, closed envelope of the set $G_1 = \{x_{mn}\}$ ($m, n = 1, 2, \cdots$). As every closed linear subspace of E is weakly closed, we have $\bar{G}_1 \subset E_1$ and therefore $\bar{G}_1$ is separable. Considering now G_1 and $\bar{G}_1$ as sets in E^{**} and applying to them theorem 21 we conclude that $\bar{G}_1$ is a weakly continuous image of a compact metric set and consequently is weakly compact and weakly closed. Hence the sequence $\{x_n\}_1^\infty \subset \bar{G}_1$ contains a subsequence $\{x_{n_\nu}\}$ converging weakly to an element $x_0 \in \bar{G}_1 \subset \bar{G}$, which proves our theorem.

THEOREM 24. *If a set $G \subset E$ is weakly compact in E, so is the convex envelope K of G.*

PROOF. Let $\{x_n\}_1^\infty \subset K$. Then every x_n ($n = 1, 2, \cdots$) is a limit of a sequence of points z of the form $z = \sum \mu_i y_i$, where $y_i \in G$, $\mu_i > 0$ ($i = 1, 2, \cdots$) and $\sum \mu_i = 1$. Therefore there exists a countable set $G_1 \subset G$ such that its convex envelope contains the sequence $\{x_n\}_1^\infty$. The linear, closed envelope E_1 of G_1 is separable and consequently so is the set $\bar{G}_1 \subset E_1$. If we show that the convex envelope of $\bar{G}_1$ is weakly compact in E, then we can state that the sequence $\{x_n\}_1^\infty \subset \bar{G}_1$ contains a subsequence $\{x_{n_\nu}\}$ converging weakly to an element of E, which, consequently, proves our theorem. Thus we see that to prove our theorem it is sufficient to prove it in the special case when E is separable and the set $G \subset E$ is weakly closed.

Consider the set $G^{**} \subset E^{**}$ of all elements $x^{**} \in E^{**}$ corresponding to the elements $x \in G$. The set G being supposed weakly compact and separable, then so is the set G^{**} (considered as a set of functionals defined over E^*). Therefore we can apply to G^{**} theorem 22 and conclude that H, the regularly convex envelope of G^{**}, is weakly compact and weakly closed. Thus to show that the convex envelope of G is weakly compact in E it is sufficient to verify that every element of H is generated by an element of E.

Let $X_0 \in H$; then by theorem I $X_0 = \mathfrak{p}_0(G^{**})$, where $\mathfrak{p}_0 \in \mathfrak{P}_{G^{**}}$. Consequently, for every $f \in E^*$, $X_0(f) = \mathfrak{p}_0(\varphi_f)$, where $\varphi_f(x^{**}) = x^{**}(f) = f(x)$ ($x^{**} \in G^{**}$). The space E being separable, in order to prove that there exists an $x_0 \in E$ such that $X_0 = x_0^{**}$, i.e. such that $X_0(f) = f(x_0)$ ($f \in E^*$), it is sufficient to show (see Banach [2], p. 131) that the functional $X_0(f)$ ($f \in E^*$) is weakly continuous.

Let a sequence $\{f_n\}_1^\infty \in E^*$ converge weakly to an element $f_0 \in E^*$. Then

$$\lim_{n\to\infty} \varphi_{f_n}(X) = \varphi_{f_0}(X) \qquad (X \in G^{**}). \tag{56}$$

Let L be the linear space of all weakly continuous functions $\varphi(X)$ ($X \in G^{**}$), which we shall consider as a subspace of $Q_{G^{**}}$ (see Chap. I, §1). According to theorem 21 we may introduce in G^{**} such a metric that G^{**} becomes a compact

metric set and L the set of all continuous functions defined over this set. Evidently $\varphi_f \in L$ $(f \in E^*)$ and

$$\| \varphi_f \| = \operatorname*{Sup}_{X \in G^{**}} | \varphi_f(X) | \leqq C |f|,$$

where C depends only on G^{**}.

Consequently, the relation (56) means (see Banach, [2] p. 224) that the sequence $\{\varphi_{f_n}\}$ converges weakly to φ_{f_0} in the sense of weak convergence of elements of L and, consequently, of elements of L. But then

$$\lim_{n \to \infty} \mathfrak{p}_0(\varphi_{f_n}) = \mathfrak{p}_0(\varphi_{f_0}),$$

i.e. $X_0(f_n) \to X_0(f_0)$, which completes the proof of our theorem.

COROLLARY. *If a set $G \subset E$ is weakly compact in E, then its convex closed envelope K is weakly compact in itself.*

In fact, K being convex and closed is also weakly closed.[16] Consequently K is the weak closure of the convex envelope K_1 of G. By theorem 24, K_1 is weakly compact in E and consequently K is weakly closed in itself according to theorem 23.

3. Invariant points of transformations. Generalizing some investigations of G. D. Birkhoff and O. D. Kellogg, J. Schauder among others has established the following theorem:

A. *Let $F(x)$ be a continuous operation, which transforms a convex, closed set $H \subset E$ into its compact part. Then $F(x)$ has a fixed point, i.e. there exists such an $x_0 \in H$ that $F(x_0) = x_0$.*

Using this theorem, J. Schauder obtains, in addition, the theorem

B. *Let E be a separable Banach space and let $H \subset E$ be a convex, weakly compact and weakly closed set. If a weakly continuous operation $F(x)$ defined over H transforms the set H into its part, then $F(x)$ has a fixed point $x_0 \in H$ $(F(x_0) = x_0)$.*

The results of the preceding paragraph permit one to obtain from J. Schauder's propositions A and B two generalizations of proposition B.

THEOREM 25. *Let $H \subset E$ be a convex, closed set. If a weakly continuous operation $F(x)$ defined over H transforms H into its separable and weakly compact part in E, then $F(x)$ has a fixed point.*

PROOF. Let H_1 be the smallest convex, closed set containing $G = F(H)$. By the corollary to theorem 24 the set H_1 is weakly compact and weakly closed. Furthermore, H_1, being evidently separable, belongs to a separable space E_1. On the other hand we have, obviously, $F(H_1) \subset F(H) \subset H_1$. To prove our theorem it remains to apply J. Schauder's proposition B to the operation $F(x)$ considered over H_1.

THEOREM 26. *Let $K \subset E^*$ be a convex, separable and weakly closed set. If a*

[16] S. Mazur [10], p. 80.

weakly continuous operation $F(f)$ transforms K into its weakly compact part G in E^*, then $F(f)$ has a fixed point.

Proof. Let K_1 be the smallest convex and weakly closed set containing the set $G = F(K)$. By theorem 22 K_1 is weakly compact, weakly closed, and separable. Denote by F the linear envelope of the set K_1. In virtue of Lemma I of Chap. I, §4, we can built a linear, separable space $E_1 \subset E$ such that $F_1 \subset E_1^*$. As E_1 is separable we can introduce a new norm in E_1^* such that every sphere $\sigma(\theta, r) \epsilon E_1$ and consequently the set K_1 becomes by the new norm a compact metric set in which the convergence in the new sense coincides with the original weak convergence (see the proof of the Lemma I, §1). After this our theorem becomes an immediate consequence of theorem A of J. Schauder.

REFERENCES

(1) N. Achyeser and M. Krein, *On some questions in the moment problem,* (in Russian) Kharkoff, 1939.

(2) S. Banach, *Théorie des opérations linéaires,* Warszawa, 1932.

(3) V. Gantmacher and V. Šmulian, *Sur les espaces linéaires dont la sphère unitaire est faiblement compacte,* C. R. Acad. Sci. URSS, vol. 17, N3 (1937), pp. 91–94.

(4) F. Hausdorf, *Zur theorie der linearen metrischen Räume,* Journal für reine und angew. Math., vol. 167 (1932), pp. 294–311.

(5) L. Kantorovitch, *On the moment problem for a finite interval,* C. R. Acad. Sci. URSS, vol. 14 (1937), pp. 531–537.

(6) M. Krein, *Sur les fonctionelles positives additives dans les espaces linéaires normés,* Comm. de la Soc. Math. de Kharkoff, sér. 4, vol. XIV, (1937), pp. 227–237.

(7) M. Krein, *Sur quelques questions de la géometrie des ensembles convexes situés dans un espace linéaire normé et complet,* C. R. Acad. Sci. URSS, vol. 14 (1937), pp. 5–7.

(8) L. Lusternik, *Principal notions of Functional Analysis* (in Russian), Advances of Mathematics, vol. I (1936), pp. 77–140.

(9) S. Mazur, *Über convexe Mengen in linearen normierten Räumen,* Studia Mathematica, vol. 4 (1933), pp. 70–84.

(10) S. Mazur, *Über die kleinste konvexe Mengen, die eine gegebene kompacte Menge enthält,* Studia Mathematica, vol. 2 (1930), pp. 7–9.

(11) B. Pettis, *A note on regular Banach spaces,* Bull. Amer. Math. Soc., vol. 44, (1938), pp. 420–428.

(12) M. Riesz, *Sur le problème des moments (Troisieme note),* Arkiv for Mathematik, Astronomi och Fysik, vol. 17 (1923), pp. 1–52.

(13) J. Schauder, *Der Fixpunktsatz in Funktionalräumen,* Studia Mathematica, Vol. 2 (1930), pp. 171–180.

(14) I. J. Schoenberg, *Convex domains and linear combinations of continuous functions,* Bull. Amer. Math. Soc., vol. 39 (1933), pp. 273–280.

(15) V. Šmulian, *Sur les ensembles régulièrement fermés et faiblement compacts dans l'espace du type* (B). C. R. Acad. Sci. URSS, vol. 18 (1938), pp. 405–407.

(16) V. Šmulian, *On the inclusion principle in the space of* (B) *type,* Recueil Mathématique, Moscou, 5 (1939) (in press).

(17) V. Šmulian, (W. Chmoulyan), *Sur les ensembles faiblement compacts dans les espaces linéaires normés,* Comm. de la Soc. Math. de Kharkoff, sér. 4, t. XIV, 1937, pp. 239–242.

Institute of Mathematics,
Kharkoff, USRS.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

SOME PROBLEMS IN THE THEORY OF SINGULAR INTEGRAL EQUATIONS

BY W. J. TRJITZINSKY

(Received June 1, 1939)

1. Introduction

Our present object is to investigate the solutions of the singular integral equations

$$\varphi(x) - \lambda \int_0^1 K(x, y)\varphi(y)\, dy = f(x), \tag{1.1}$$

$$\varphi(x) - \lambda \int_0^1 K(x, y)\varphi(y)\, dy = 0 \tag{1.1a}$$

for real values of the parameter λ, *under the assumption that*

$$f(x) \subset L_2 \qquad (x \text{ on } (0, 1)) \tag{1.1b}$$

and that the real symmetric kernel $K(x)$ *is measurable.*[1]

We shall also consider the corresponding integral equation of the first kind.

In the sequel various additional hypotheses with respect to $K(x, y)$ will be made. The equations (1.1), (1.1a) will be *singular*, inasmuch as it will be not required that

$$\int_0^1 \int_0^1 K^2(x, y)\, dx\, dy \tag{1.2}$$

should exist. As is well known, the essential features of the Fredholm theory of integral equations will hold for (1.1), (1.1a) when the integral (1.2) exists (in the Lebesgue sense).

For λ non-real, as remarked by *T. Carleman*,[2] the situation is simpler than for λ real. On the other hand, applications of integral equations to dynamics and quantum mechanics are of particular importance for λ real. This fact explains our present purpose. For λ non-real Carleman developed a highly important theory of singular integral equations. This theory has been extended again for λ non-real, in a work by *W. J. Trjitzinsky.*[3]

We shall denote by S the square

$$0 \leq x, \qquad y \leq 1. \tag{1.3}$$

---

[1] Measurability in this paper is in the Lebesgue sense.

[2] T. Carleman, *Sur les équations intégrales singulières à noyau réel et symétrique* [Uppsala, 1923]; in the sequel referred to as (C).

[3] W. J. Trijitzinsky, *General theory of singular linear integral equations with real kernels* [Transactions Amer. Math. Soc. (1939), 202–279]. In this paper kernels are considered which, while not necessarily of the type considered in (C), are repeated limits of kernels of Carleman type.

For n a positive integer let E_n be the part of S in which

$$| K(x, y) | \leqq n.$$

One may define[4] $K_n(x, y)$ as follows:

$$\begin{aligned} K_n(x, y) &= K(x, y) && \text{(in } E_n\text{)}, \\ K_n(x, y) &= n && \text{(when } K(x, y) > n\text{)}, \\ K_n(x, y) &= -n && \text{(when } K(x, y) < -n\text{)}. \end{aligned} \tag{1.4}$$

Clearly

$$| K_n(x, y) | \leqq | K_{n+1}(x, y) | \leqq | K(x, y) |, \quad \lim_n K_n(x, y) = K(x, y). \tag{1.4a}$$

In the sequel the kernels under consideration will be of one of the following three types.

TYPE 1.A. The integral

$$\int_0^1 K^2(x, y)\, dy \tag{1.5}$$

exists for $x \neq \xi_\nu$ ($\nu = 1, 2, \cdots$), where the ξ_ν possess a finite number of limiting points, included amongst the points

$$\eta_1, \cdots \eta_m. \tag{1.5a}$$

The set (ξ_ν) will be taken closed. It is supposed that if one defines symmetric $K^{(\delta)}(x, y)$ $(\delta > 0)$ by the relations

$$\begin{aligned} K^{(\delta)}(x, y) &= 0 && (0 \leqq y \leqq 1; | x - \eta_\nu | < \delta; \nu = 1, \cdots m), \\ K^{(\delta)}(x, y) &= K(x, y) && \text{(elsewhere)}, \end{aligned} \tag{1.5b}$$

then the integral

$$\int_0^1 \int_0^1 K^{(\delta)2}(x, y)\, dx\, dy \tag{1.6}$$

exists for all $\delta > 0$. Moreover,

$$\lim_{x_1 \to x_2} \int_0^1 [K(x_1, y) - K(x_2, y)]^2\, dy = 0 \qquad (x_1, x_2 \neq \xi_\nu; \nu = 1, 2, \cdots).\text{[5]} \tag{1.7}$$

TYPE 1.B. The integral

$$\int_0^1 K^2(x, y)\, dy \tag{1.8}$$

exists for almost all x (on $(0, 1)$).[6]

---

[4] Cf. T. Carleman, *La théorie des équations intégrales singulières et les applications* [Annales de l'Institut H. Poincaré *1* (1931), 401–430]; in the sequel referred to as (C_1).

[5] A comprehensive theory, for λ non-real, for kernels of Type 1. A has been developed by Carleman in the greater part of (C).

[6] Kernels of Type 1.B have been involved in a number of developments due to Carleman; cf. (C_1) as well as a number of other papers by Carleman, giving applications to problems of dynamics and quantum mechanics.

Type 1.C. Nothing is added to the previously made hypothesis (to the effect that the symmetric kernel $K(x, y)$ be measurable) except the following indirect condition.

Associated with $K(x, y)$ there exists a linear operator $L_x(\xi \mid h(x))$ (ξ a parameter) such that the following five conditions hold:

$$L_x(\xi \mid K(x, y)) \subset L_2 \quad \text{(in } y\text{);} \tag{1.9}$$

$$| L_x(\xi \mid K_n(x, y)) | < \gamma(\xi \mid y) \subset L_2 \quad (\text{in } y; \gamma(\xi \mid y) \text{ independent of } n), \tag{1.9a}$$

where $K_n(x, y)$ is from (1.4);

$$\lim_n L_x(\xi \mid K_n(x, y)) = L_x(\xi \mid K(x, y)); \tag{1.9b}$$

$$\lim_n L_x(\xi \mid f_n(x)) = L_x(\xi \mid f(x)), \quad \text{whenever} \quad f_n(x) \subset L_2 \quad \text{and} \quad f_n(x) \to f(x) (\text{weakly});^7 \tag{1.9c}$$

$$\int_0^1 L_x(\xi \mid K_n(x, y))\varphi(y)\, dy = L_x\left(\xi \,\middle|\, \int_0^1 K_n(x, y)\varphi(y)\, dy\right) \tag{1.9d}$$

for all $\varphi(x) \subset L_2$.[8]

It is clear that the types of kernels, 1.A, 1.B, 1.C, as described above, are of increasing degrees of generality, in the order stated.

For kernels which are merely measurable to (1.1), (1.1a) there will correspond the equations

$$\varphi(x) - \lambda \int_0^1 K_n(x, y)\varphi(y)\, dy = f(x), \tag{1.10}$$

$$\varphi(x) - \lambda \int_0^1 K_n(x, y)\varphi(y)\, dy = 0 \quad (K_n(x, y) \text{ from (1.4)}). \tag{1.10a}$$

On the other hand, for kernels of Type 1.A, corresponding to (1.1) and (1.1a) we shall also have

$$\varphi(x) - \lambda \int_0^1 K^{(\delta)}(x, y)\varphi(y)\, dy = f(x), \tag{1.11}$$

$$\varphi(x) - \lambda \int_0^1 K^{(\delta)}(x, y)\varphi(y)\, dy = 0 \quad (\delta > 0), \tag{1.11a}$$

where $K^{(\delta)}(x, y)$ is from (1.5b).

Finally, of importance for kernels of Type 1.C will be the equations

$$L_x(\xi \mid \varphi(x)) - \lambda \int_0^1 L_x(\xi \mid K_n(x, y))\varphi(y)\, dy = L_x(\xi \mid f(x)), \tag{1.12}$$

---

[7] $f_\nu(x) \to f(x)$ weakly, if $\lim_n \int_0^x f_n(x)\, dx = \int_0^x f(x)\, dx$.

[8] Conditions (1.9), $\cdots$, (1.9d) are analogous to those employed in (C; p. 138) in an investigation (for λ non-real) of equations whose kernels belong to a variety somewhat less general than that involved in Type 1.C.

$$L_x(\xi \mid \varphi(x)) - \lambda \int_0^1 L_x(\xi \mid K(x, y))\varphi(y)\, dy = L_x(\xi \mid f(x)), \tag{1.12a}$$

where L is an operator satisfying (1.9), $\cdots$ (1.9d).

2. A lemma valid for all equations (1.1)

Suppose $K_n(x, y)$ is defined by (1.4). Inasmuch as $K_n(x, y)$ is measurable and

$$| K_n(x, y) | \leqq n,$$

it is inferred that the corresponding homogeneous integral equation (1.10a) possesses a set of real characteristic values

$$\lambda_{n,\nu} \qquad (\nu = 1, 2, \cdots), \tag{2.1}$$

containing at least one value; moreover, for n fixed, the set (2.1) will have no finite limiting values. Let the corresponding characteristic functions, arranged as an ortho-normal set, be

$$\varphi_{n,\nu}(x) \qquad (\nu = 1, 2, \cdots; \varphi_{n,\nu}(x) \subset L_2); \tag{2.1a}$$

thus,

$$\varphi_{n,\nu}(x) - \lambda_{n,\nu} \int_0^1 K_n(x, y)\varphi_{n,\nu}(y)\, dy = 0. \tag{2.1b}$$

Form an infinite subsequence

$$n_j \qquad (j = 1, 2, \cdots; n_1 < n_2 < \cdots) \tag{2.2}$$

of the set (n) of positive integers. *Let G be the set of points represented by the numbers*

$$\lambda_{n_j,\nu} \qquad (j, \nu = 1, 2, \cdots). \tag{2.2a}$$

The closure[9] $\bar{G}$ of G may consist of all the points of the real axis in the complex λ-plane.

Throughout this paper we proceed under the hypothesis that at least for some subsequence (2.2) the above set $\bar{G}$ does not consist of all the points of the real λ-axis. $\bar{G}$ being closed, the set $C\bar{G}$[10] will then be open and, thus, will contain an interval

$$\Delta = \{| \lambda_0 - \lambda | < \omega\}. \tag{2.3}$$

Whenever $K(x, y)$ is of Type 1.A, we form an infinite subsequence

$$\delta_j \qquad (\delta_j > 0; j = 1, 2, \cdots; \lim_j \delta_j = 0) \tag{2.4}$$

and designate by G the set of points represented by the numbers

$$\lambda_\nu^{(\delta_j)} \qquad (j, \nu = 1, 2, \cdots), \tag{2.4a}$$

[9] $\bar{G} = G +$ limiting points of G.
[10] $C\bar{G} =$ real λ-axis $- \bar{G}$.

where the $\lambda_\nu^{(\delta_j)}$ $[\nu = 1, 2, \cdots]$ form the set of characteristic values for the equation (1.11a) (with $\delta = \delta_j$).[11] The corresponding characteristic functions (arranged as an ortho-normal set)

$$\varphi_\nu^{(\delta_j)}(x) \qquad (\nu = 1, 2, \cdots; \varphi_\nu^{(\delta_j)}(x) \subset L_2) \tag{2.4b}$$

will satisfy

$$\varphi_\nu^{(\delta_j)}(x) - \lambda_\nu^{(\delta_j)} \int_0^1 K^{(\delta_j)}(x, y)\varphi_\nu^{(\delta_j)}(y)\, dy = 0. \tag{2.4c}$$

Whenever it is explicitly stated that $K(x, y)$ is of Type 1.A the supposition will be made that at least for some subsequence (2.4) the closure $\bar{G}$ of the set G (referred to in connection with (2.4a)) will not consist of the whole of the real λ-axis; $C\bar{G}$ will certainly contain an open interval

$$\Delta = \{|\lambda_0 - \lambda| < \omega\}. \tag{2.5}$$

DEFINITION 2.1. *With an interval Δ, defined by (2.3) or (2.5) (as the case may be), let d denote a sub interval,*

$$d = \{|\lambda_0 - \lambda| \leqq h\} \qquad (0 < h < \omega), \tag{2.6}$$

of Δ. We then let S_d designate the set of points consisting of the interval d and of the points of the λ-plane corresponding to all the non-real values λ.

For λ on d

$$|\lambda - \lambda_{n_j,\nu}| \geqq \omega - h \qquad [>0; j, \nu = 1, 2, \cdots]. \tag{2.7}$$

When $K(x, y)$ is of Type 1.A then, for λ on d,

$$|\lambda - \lambda_\nu^{(\delta_j)}| \geqq \omega - h. \tag{2.7a}$$

Let λ be fixed on d and let $\varphi_n(x) \subset L_2$ $(n = n_j)$ be the solution of (1.10),

$$\varphi_n(x) - \lambda \int_0^1 K_n(x, y)\varphi_n(y)\, dy = f(x), \tag{2.8}$$

for this value λ. We multiply (2.8) by $\bar{\varphi}_n(x)\, dx$ and integrate, obtaining[12]

$$\int_0^1 |\varphi_n(x)|^2\, dx = \lambda \int_0^1\int_0^1 K_n(x, y)\varphi_n(y)\bar{\varphi}_n(x)\, dx\, dy + \int_0^1 f(x)\bar{\varphi}_n(x)\, dx. \tag{2.8a}$$

Before establishing an inequality for the double integral in (2.8a) we shall consider the function

$$\tau_n(x) = \int_0^1 K_n(x, y)\varphi_n(y)\, dy. \tag{2.9}$$

---

[11] Note the statement with respect to (1.6).

[12] The double integral in (2.8a) is real since $K_n(x, y)$ is symmetric; thus, the imaginary part of the integral last displayed in (2.8a) is zero.

In view of a known theorem, $\tau_n(x)$ can be expanded in a convergent series in terms of the functions (2.1a). Thus

$$\tau_n(x) = \sum_{\nu=1}^{\infty} \frac{1}{\lambda_{n,\nu}} \varphi_{n,\nu}\bar{\varphi}_{n,\nu}(x), \tag{2.9a}$$

where

$$\varphi_{n,\nu} = \int_0^1 \varphi_n(y)\varphi_{n,\nu}(y)\,dy; \qquad \bar{\varphi}_{n,\nu} = \int_0^1 \bar{\varphi}_n(y)\varphi_{n,\nu}(y)\,dy. \tag{2.9b}$$

Designate by w_n the double integral in (2.8a); by (2.9), (2.9a), (2.9b), integration term by term being permissible, we have

$$\begin{aligned} w_n = \int_0^1 \tau_n(x)\bar{\varphi}_n(x)\,dx &= \sum_{\nu=1}^{\infty} \frac{1}{\lambda_{n,\nu}} \varphi_{n,\nu} \int_0^1 \bar{\varphi}_n(x)\varphi_{n,\nu}(x)\,dx \\ &= \sum_{\nu=1}^{\infty} \left(\frac{\varphi_{n,\nu}}{\lambda_{n,\nu}}\right) \bar{\varphi}_{n,\nu}. \end{aligned} \tag{2.10}$$

Applying the Schwarzian inequality to the series last displayed, one obtains

$$|w_n|^2 \leqq \sum_{\nu=1}^{\infty} \left|\frac{\varphi_{n,\nu}}{\lambda_{n,\nu}}\right|^2 \sum_{\nu=1}^{\infty} |\varphi_{n,\nu}|^2.$$

In view of (2.9b) and of Bessel's inequality

$$|w_n|^2 \leqq \int_0^1 |\varphi_n(x)|^2\,dx \sum_{\nu=1}^{\infty} \left|\frac{\varphi_{n,\nu}}{\lambda_{n,\nu}}\right|^2 \qquad (n = n_j). \tag{2.10a}$$

On writing

$$f_{n,\nu} = \int_0^1 f(x)\varphi_{n,\nu}(x)\,dx \tag{2.10b}$$

we have a known relation

$$\frac{\varphi_{n,\nu}}{\lambda_{n,\nu}} = \frac{f_{n,\nu}}{\lambda_{n,\nu} - \lambda} \qquad [\lambda \text{ on } d;\ n = n_j;\ j, \nu = 1, 2, \cdots].$$

Consequently, in view of (2.10a), of (2.7) and of Bessel's inequality (applicable because of (2.10b))

$$\begin{aligned} |w_n|^2 &\leqq \int_0^1 |\varphi_n(x)|^2\,dx \sum_{\nu=1}^{\infty} \left|\frac{f_{n,\nu}}{\lambda_{n,\nu} - \lambda}\right|^2 \\ &\leqq \frac{1}{(\omega - h)^2} \int_0^1 |\varphi_n(x)|^2\,dx \sum_{\nu=1}^{\infty} |f_{n,\nu}|^2 \\ &\leqq \frac{1}{(\omega - h)^2} \int_0^1 |\varphi_n(x)|^2\,dx \int_0^1 |f(x)|^2\,dx = M_n^2. \end{aligned} \tag{2.11}$$

Application of (2.11) and of the Schwarzian inequality to (2.8a) will yield

$$\int_0^1 |\varphi_n(x)|^2\,dx \leqq |\lambda| M_n + \left[\int_0^1 |f(x)|^2\,dx\right]^{\frac{1}{2}}\left[\int_0^1 |\varphi_n(x)|^2\,dx\right]^{\frac{1}{2}}$$

$$\leqq \left[1 + \frac{|\lambda|}{\omega - h}\right]\left[\int_0^1 |f(x)|^2\,dx\right]^{\frac{1}{2}}\left[\int_0^1 |\varphi_n(x)|^2\,dx\right]^{\frac{1}{2}}.$$

Accordingly, for $j = 1, 2, \cdots$,

$$\int_0^1 |\varphi_n(x)|^2\,dx \leqq \left[1 + \frac{|\lambda|}{\omega - h}\right]^2 \int_0^1 |f(x)|^2\,dx \qquad (\lambda \text{ on } d;\ n = n_j). \tag{2.12}$$

An analogous inequality will hold when $K(x, y)$ is of type 1.A.

Hence the following Lemma has been established.

LEMMA 2.1. *Let symmetric $K(x, y)$ be merely measurable. Let d be an interval (supposed to exist) as specified in Definition 2.1. With $\varphi_{n_j}(x) \subset L_2$ designating solution of (1.10) for a value λ fixed on d, we shall have the inequality (2.12) satisfied as stated.*[13] *If $K(x, y)$ is of Type 1.A and d is an interval (supposed to exist for some subsequence δ_j; cf. (2.4)) as described in Definition (2.1), one will have*

$$\int_0^1 |\varphi^{(\delta_j)}(x)|^2\,dx \leqq \left[1 + \frac{|\lambda|}{\omega - h}\right]^2 \int_0^1 |f(x)|^2\,dx \qquad [\lambda \text{ on } d;\ j = 1, 2, \ldots], \tag{2.12a}$$

where $\varphi^{(\delta_j)}(x) \subset L_2$ is solution of (1.11) (with $\delta = \delta_j$) for a value λ fixed on d.[14]

3. Existence of solutions for non-homogeneous equations

We recall a theorem of *F. Riesz* according to which inequalities

$$\int_0^1 |h_\nu(x)|^2\,dx \leqq M \qquad (\nu = 1, 2, \cdots;\ M \text{ independent of } \nu) \tag{3.1}$$

imply existence of a subsequence $\{h_{\nu_j}(x)\}$ $(\nu_1 < \nu_2 < \cdots)$ and of a function $h(x)$ so that

$$\lim_j h_{\nu_j}(x) = h(x) \qquad (\text{weakly}),^7 \tag{3.1a}$$

while

$$\int_0^1 |h(x)|^2\,dx \leqq M. \tag{3.1b}$$

With symmetric $K(x, y)$ merely measurable in λ fixed on d, the above theorem, together with inequalities (2.12) of Lemma 2.1, will imply that there exists a subsequence $\{\varphi_{m_j}(x)\}$ of $\{\varphi_{n_j}(x)\}$ and a function $\varphi(x) \subset L_2$ so that

[13] The $n_j (j = 1, 2, \cdots)$ in (2.12) are the numbers (2.2), involved in (2.2a) in connection with the definition of G; also see the italics subsequent to (2.2a).

[14] The inequalities (2.12), (2.12a) are analogous to certain inequalities obtained in (C) for λ non-real; it is to be noted that in the latter case selection of subsequences (n_j) or (δ_j) is not necessary. The formulas stated in (C) for λ non-real are not valid for λ real.

$$\lim_{j} \varphi_{m_j}(x) = \varphi(x) \quad \text{(weakly)} \tag{3.2}$$

and

$$\int_0^1 |\varphi(x)|^2 dx \leqq \left[1 + \frac{|\lambda|}{\omega - h}\right]^2 \int_0^1 |f(x)|^2 dx = q^2. \tag{3.2a}$$

Unless additional hypotheses are introduced it is generally out of question to examine the possibility of $\varphi(x)$ representing a solution of (1.1). We shall accordingly examine the situation when $K(x, y)$ belongs to one of the three Types 1.A, 1.B, 1.C.

When $K(x, y)$ is of Type 1.A application of (2.12a) enables one to infer existence of a subsequence $\{\varphi^{(\delta_j')}(x)\}$ of $\{\varphi^{(\delta_j)}(x)\}$ such that there exists a function $\varphi(x)$ for which

$$\lim_{j} \varphi^{(\delta_j')}(x) = \varphi(x) \quad \text{(weakly)}, \tag{3.3}$$

$$\int_0^1 |\varphi(x)|^2 dx \leqq q^2 \quad [q \text{ from (3.2a)}; \lambda \text{ fixed on } d]. \tag{3.3a}$$

Now, by definition of $\varphi^{(\delta)}(x)$ $(\delta > 0)$

$$\varphi^{(\delta)}(x) = \lambda \int_0^1 K^{(\delta)}(x, y)\varphi^{(\delta)}(y)\, dy + f(x) \quad (\delta = \delta_j'). \tag{3.4}$$

According to a known theorem,[15] if $f_\nu(x)$ and $g_\nu(x) \subset L_2$ and

$$\int_0^1 |f_\nu(x)|^2 dx < M \tag{3.5}$$

and $f_\nu(x) \to f(x)$ weakly, while $\lim_\nu g_\nu(x) = g(x)$ (almost everywhere) and $|g_\nu(x)| < \gamma(x) \subset L_2$, then

$$\lim_\nu \int_0^1 f_\nu(x)g_\nu(x)\, dx = \int_0^1 f(x)g(x)\, dx. \tag{3.5a}$$

In view of the convergence of the integral (1.5) and since by definition (1.5b)

$$|K^{(\delta_j')}(x, y)| \leqq |K(x, y)| \subset L_2 \quad \text{(in } y\text{)},$$

on taking account of (2.12a) (with δ_j replaced by δ_j') and of (3.3), in consequence of the result [(3.5), (3.5a)] it is inferred that

$$\lim_j \int_0^1 K^{(\delta_j')}(x, y)\varphi^{(\delta_j')}(y)\, dy = \int_0^1 K(x, y)\varphi(y)\, dy \tag{3.6}$$

$[\lambda$ fixed on d; $x \neq \xi_\nu$ (cf. statement in connection with (1.5a))$]$.

---

[15] (C; 132–133).

By virtue of (3.6) and (3.4) it is deduced that the limit

$$\lim_i \varphi^{(\delta'_i)}(x) = \psi(x) = \lambda \int_0^1 K(x, y)\varphi(y)\, dy + f(x) \tag{3.6a}$$

exists (for $x \neq \xi_\nu$; $\nu = 1, 2, \cdots$). By (3.3)

$$\lim_i \int_0^x \varphi^{(\delta'_i)}(x)\, dx = \int_0^x \varphi(x)\, dx = \text{absolutely continuous function.} \tag{3.6b}$$

On the other hand,

$$\lim_i \int_0^x \varphi^{(\delta'_i)}(x)\, dx = \int_0^x \lim_i \varphi^{(\delta'_i)}(x)\, dx = \int_0^x \varphi(x)\, dx, \tag{3.6c}$$

inasmuch as the passage to the limit under the integral sign in (3.6c) is justifiable by virtue of (3.3), of (2.12a) (with δ_i replaced by δ'_i) and of theorem [(3.5), (3.5a)]. By (3.6a) and (3.6c)

$$\int_0^x \psi(x)\, dx = \int_0^x \varphi(x)\, dx$$

and, accordingly

$$\psi(x) = \varphi(x) \qquad \text{(almost everywhere on (0. 1)).}$$

This, in view of (3.6a), implies that *the function* $\varphi(x)$, *involved in* (3.3), *is a solution* (*almost everywhere*) *of the equation*

$$\varphi(x) = \lambda \int_0^1 K(x, y)\varphi(y)\, dy + f(x) \qquad (K(x, y) \text{ of Type 1.A}) \tag{3.7}$$

for λ *fixed on the interval d.* More precise results for kernels of Type 1.A will be given in Theorem 3.2.

Suppose now that $K(x, y)$ is of Type 1.B. In place of $\varphi^{(\delta'_i)}(x)$ we now have $\varphi_{m_j}(x)$ (cf. (3.2), (3.2a)). Repeating the argument from (3.3) to (3.7), with $K^{(\delta'_i)}(x, y)$ replaced by $K_{m_j}(x, y)$ and noting that

$$| K_{m_j}(x, y) | \leqq | K(x, y) | \subset L_2 \qquad \text{(in } y\text{; almost all } x\text{)}$$

(cf. (1.4a) and (1.8)), it is observed that *the function* $\varphi(x)$ *of* (3.2) *satisfies for almost all x* (*on* (0, 1) *the equation*

$$\varphi(x) = \lambda \int_0^1 K(x, y)\varphi(y)\, dy + f(x) \qquad (K(x, y) \textit{ of Type } 1.\text{B}) \tag{3.7a}$$

for λ *fixed on the interval d.*

If $K(x, y)$ is of type 1.C there exists a linear operator L satisfying, with respect to $K(x, y)$, the conditions (1.9), $\cdots$ (1.9d). Let $\{\varphi_{m_j}(x)\}$ be the sequence, involved in (3.2), for this kernel $K(x, y)$ and for λ fixed on d. Thus

$$\varphi_{m_j}(x) = \lambda \int_0^1 K_{m_j}(x, y)\varphi_{m_j}(y)\, dy + f(x), \tag{3.8}$$

(3.8a) $\int_0^1 |\varphi_{m_j}(x)|^2 dx \leqq q^2$ (cf. (3.2a)), $\lim_j \varphi_{m_j}(x) = \varphi(x)$ (weakly).

From (3.8) one obtains

$$L_x(\xi \mid \varphi_{m_j}(x)) = \lambda \int_0^1 L_x(\xi \mid K_{m_j}(x, y))\varphi_{m_j}(y)\, dy + L_x(\xi \mid f(x)). \tag{3.9}$$

By (1.9c) and (3.8a)

$$\lim_j L_x(\xi \mid \varphi_{m_j}(x)) = L_x(\xi \mid \varphi(x)). \tag{3.9a}$$

On noting that (1.9b) and (1.9a) (with $n = m_j$) hold, in consequence of (3.8a) it is inferred that theorem [(3.5), (3.5a)] is applicable, yielding the result

$$\lim_j \int_0^1 L_x(\xi \mid K_{m_j}(x, y))\varphi_{m_j}(y)\, dy = \int_0^1 L_x(\xi \mid K(x, y))\varphi(y)\, dy. \tag{3.9b}$$

In view of (3.9a) and (3.9b) from (3.9) it follows that $\varphi(x)$ *satisfies the equation* (1.12a). We sum the above results in the following theorem.

THEOREM 3.1. *It is recalled that when $K(x, y)$ is of Type 1.B or of Type 1.C there exists, by hypothesis, a sequence (n_j) (2.2), associated with which there is on hand an interval Δ (cf. (2.3) and the text from (2.2) to (2.3)). Similarly, when $K(x, y)$ is of Type 1.A there exists, by hypothesis, a sequence (δ_j) (2.4) and an associated interval Δ (cf. (2.5) and the text from (2.4) to (2.5)). In either case, let d denote a sub-interval (2.6) of Δ and let λ denote a fixed value in d.*

In the case of Type 1.A, for a subsequence $\{\varphi^{(\delta_j')}(x)\}$ of $\{\varphi^{(\delta_j)}(x)\}$ we have (3.3) and the function $\varphi(x)$ involved in (3.3) will be a solution for the value λ in question and for almost all x on (0, 1) of the equation (1.1).

In the case of Type 1.B, for a subsequence $\{\varphi_{m_j}(x)\}$ of $\{\varphi_{n_j}(x)\}$ we have (3.2) and $\varphi(x)$ (in (3.2)) will be a solution for the above λ and for almost all x on (0, 1) of the equation (1.1).

When $K(x, y)$ is of Type 1.C so that an operator L, satisfying (1.9), $\cdots$ (1.9d) is associated with $K(x, y)$, we have (3.2) for a subsequence $\{\varphi_{m_j}(x)\}$ of $\{\varphi_{n_j}(x)\}$; the function $\varphi(x)$ involved in (3.2) will be a solution, for the value λ in question, of the equation (1.12).

In all cases the solution $\varphi(x)$ (for the above λ) will satisfy the inequality

$$\int_0^1 |\varphi(x)|^2 dx \leqq q^2 \qquad [q^2 \text{ from (3.2a)}].$$

With symmetric $K(x, y)$ merely measurable we note the relation

$$\varphi_{m_j}(x) = \lambda \int_0^1 K_{m_j}(x, y)\varphi_{m_j}(y)\, dy + f(x). \tag{3.10}$$

When $K(x, y)$ is of Type 1.A one has

$$\varphi^{(\delta_j')}(x) = \lambda \int_0^1 K^{(\delta_j')}(x, y)\varphi^{(\delta_j')}(y)\, dy + f(x). \tag{3.10a}$$

From (3.10) and (3.10a) with the aid of the Schwarzian inequality it is established that, when $K(x, y)$ is of type 1.B,

$$(3.11) \qquad |\varphi_{m_j}(x)| \leqq \lambda_1 \left[\int_0^1 K_{m_j}^2(x, y)\,dy\right]^{\frac{1}{2}} \left[\int_0^1 |\varphi_{m_j}(y)|^2\,dy\right]^{\frac{1}{2}} + |f(x)|,$$

$$(3.11a) \qquad |\varphi_{m_j}(x) - f(x)| \leqq \lambda_1 \left[\int_0^1 K_{m_j}^2(x, y)\,dy\right]^{\frac{1}{2}} \left[\int_0^1 |\varphi_{m_j}(y)|^2\,dy\right]^{\frac{1}{2}}$$

and that, when $K(x, y)$ is of Type 1.A, we have with $\delta = \delta_j'$,

$$(3.12) \qquad |\varphi^{(\delta)}(x)| \leqq \lambda_1 \left[\int_0^1 K^{(\delta)^2}(x, y)\,dy\right]^{\frac{1}{2}} \left[\int_0^1 |\varphi^{(\delta)}(y)|^2\,dy\right]^{\frac{1}{2}} + |f(x)|,$$

$$(3.12a) \qquad |\varphi^{(\delta)}(x) - f(x)| \leqq \lambda_1 \left[\int_0^1 K^{(\delta)^2}(x, y)\,dy\right]^{\frac{1}{2}} \left[\int_0^1 |\varphi^{(\delta)}(y)|^2\,dy\right]^{\frac{1}{2}},$$

$$(3.12b) \qquad \begin{aligned} &|[\varphi^{(\delta)}(x_1) - f(x_1)] - [\varphi^{(\delta)}(x_2) - f(x_2)]|^2 \\ &\qquad \leqq \lambda_1^2 \int_0^1 [K^{(\delta)}(x_1, y) - K^{(\delta)}(x_2, y)]^2\,dy \int_0^1 |\varphi^{(\delta)}(y)|^2\,dy. \end{aligned}$$

In (3.11), $\cdots$ (3.12b)

$$(3.13) \qquad \lambda_1 = \text{upper bound of } |\lambda| \text{ for } \lambda \text{ on the interval } d.$$

By virtue of (2.12) and (2.12a), as well as in consequence of the relations

$$|K_{m_j}(x, y)| \leqq |K(x, y)| \subset L_2 \text{ (in } y), \quad |K^{(\delta_j')}(x, y)| \leqq |K(x, y)| \subset L_2 \text{ (in } y),$$

valid for the Types 1.B and 1.A, respectively, the following is deduced from (3.11), $\cdots$ (3.12b).

For the Type 1.B, almost everywhere,

$$(3.14) \qquad |\varphi_{m_j}(x)| \leqq \lambda_1 q \left[\int_0^1 K^2(x, y)\,dy\right]^{\frac{1}{2}} + |f(x)|,$$

$$(3.14a) \qquad |\varphi_{m_j}(x) - f(x)| \leqq \lambda_1 q \left[\int_0^1 K^2(x, y)\,dy\right]^{\frac{1}{2}} \qquad (q \text{ from (3.2a)}).$$

For the Type 1.A, for $x, x_1, x_2 \neq \xi_\nu$ ($\nu = 1, 2, \cdots$) and for λ on d,

$$(3.15) \qquad |\varphi^{(\delta_j')}(x)| \leqq \lambda_1 q \left[\int_0^1 K^2(x, y)\,dy\right]^{\frac{1}{2}} + |f(x)|,$$

$$(3.15a) \qquad |\varphi^{(\delta_j')}(x) - f(x)| \leqq \lambda_1 q \left[\int_0^1 K^2(x, y)\,dy\right]^{\frac{1}{2}},$$

$$(3.15b) \qquad \begin{aligned} &|[\varphi^{(\delta_j')}(x_1) - f(x_1)] - [\varphi^{(\delta_j')}(x_2) - f(x_2)]|^2 \\ &\qquad \leqq \lambda_1^2 q^2 \int_0^1 [K(x_1, y) - K(x_2, y)]^2\,dy \qquad [j \text{ sufficiently great}]. \end{aligned}$$

Thus, for kernels of Types 1.B, 1.A the sequence of functions

$$\{\varphi_{m_j}(x)\}, \qquad \{\varphi^{(\delta_j')}(x)\},$$

employed in the construction of the solutions of non-homogeneous equations, as specified in Theorem 3.1, are bounded (in absolute value) uniformly with respect to λ (λ on the interval d) and with respect to j.

When $K(x, y)$ is of Type 1.A, on noting (1.7) it is observed that the condition (3.15b) signifies that the functions of the sequence

$$\varphi^{(\delta_j')}(x) - f(x) \qquad (j = 1, 2, \cdots) \tag{3.16}$$

are continuous, for $x \neq \xi_\nu$ ($\nu = 1, 2, \cdots$; x on $(0, 1)$) uniformly with respect to λ (λ on d) and uniformly with respect to j (when x is on any closed subset of $(0, 1) - \{\xi_\nu\}$); the latter property of uniformity in accordance with the usual terminology is expressed by saying that the sequence (3.16) is equicontinuous. Thus, for a fixed value of λ (on d) a subsequence $\{\varphi^{(\delta_j')}(x)\}$ can be so selected that the limit

$$\lim_j \varphi^{(\delta_j')}(x) = \varphi(x) \tag{3.16a}$$

exists for all $x \neq \xi_\nu$ ($\nu = 1, 2, \cdots$). Now, the functions $\varphi^{(\delta_j')}(x)$ are analytic in λ for λ on S_d (Definition 2.1); on the other hand, the (δ_j') can be so selected that the limit (3.16a) exists not merely for a single value of λ (on d) but exists uniformly (with respect to λ) for λ on a preassigned denumerable set situated on the interval d;[16] these considerations lead to the conclusion that the sequence $\{\varphi^{(\delta_j')}(x)\}$ can be so selected that the limit (3.16a) exists for $x \neq \xi_\nu$ ($\nu = 1, 2, \cdots$) and for all λ in S_d; moreover, in S_d $\varphi(x)$ will be analytic in λ. Since $\varphi(x)$, as defined by (3.16a), is a particular function referred to in Theorem 3.1, it clearly is a solution $\subset L_2$ of the equation (1.1).

THEOREM 3.2. *Let $K(x, y)$ be of Type* 1.A. *The subsequence $\{\varphi^{(\delta_j')}(x)\}$, involved in Theorem* 3.1 *for the purpose of constructing a solution of the equation* (1.1), *can be so selected that the limit*

$$\lim \varphi^{(\delta_j')}(x) = \varphi(x)$$

exists for all $x \neq \xi_\nu$ ($\nu = 1, 2, \cdots$) and for all λ in S_d (Definition 2.1) and so that for every fixed $x \neq \xi_\nu$ ($\nu = 1, 2, \cdots$) $\varphi(x)$ is analytic in λ for λ in S_d. For these values of x and λ $\varphi(x)$ will constitute a solution $\subset L_2$ of the equation (1.1). *Moreover, $\varphi(x) - f(x)$ will be continuous in x (for $x \neq \xi_\nu$; $\nu = 1, 2, \cdots$) in accordance with the inequality*

$$|[\varphi(x_1) - f(x_1)] - [\varphi(x_2) - f(x_2)]|^2 \leq h \int_0^1 [K(x_1, y) - K(x_2, y)]^2\, dy$$

$$[x_1, x_2 \neq \xi_\nu \ (\nu = 1, 2, \ldots); \text{cf. } (1.6)]$$

[16] We note that the second members in (3.15), $\cdots$ (3.15b) are independent of λ. Use is made of a theorem of Vitali. Compare with (C).

where, for λ *on* d,

$$h = \lambda_1^2 q^2 \qquad \text{(cf. (3.13), (3.2a))}$$

and, for λ *non-real,* h *is a value given by Carleman.*[17]

We shall now prove the following theorem.

THEOREM 3.3. *Let* $K(x, y)$ *be of Type* 1.A *or of Type* 1.B. *For* λ *on an interval* d *a solution* $\varphi(x) \subset L_2$ *of the equation* (1.1) *is given by the formula*

$$\varphi(x) = f(x) + \lambda \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 \theta(x, y \mid \mu) f(y)\, dy \tag{3.17}$$

where $\theta(x, y \mid \mu)$ *is a spectrum of* $K(x, y)$.[18]

To establish this theorem suppose first that $K(x, y)$ is of Type 1.B. The spectrum of the equation

$$\varphi(x) = \lambda \int_0^1 K_n(x, y)\varphi(y)\, dy$$

will be

$$\theta_n(x, y \mid \lambda) = \begin{cases} \sum \varphi_{n,\nu}(x)\varphi_{n,\nu}(y) & [0 < \lambda_{n,\nu} < \lambda; \quad \text{for } \lambda > 0], \\ -\sum \varphi_{n,\nu}(x)\varphi_{n,\nu}(y) & [\lambda \leqq \lambda_{n,\nu} < 0; \quad \text{for } \lambda < 0], \end{cases}$$

while $\theta_n(x, y \mid 0) = 0$.[19] In consequence of the classical theory, which is applicable since $K_n(x, y)$ is measurable and $| K_n(x, y) | \leqq n$, the solution of (1.10) $\subset L_2$ is expressible in the form of a Stieltjes integral,

$$\begin{aligned} \varphi_n(x) &= f(x) + \lambda \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 f(y)\theta_n(x, y \mid \mu)\, dy \\ &= f(x) + \lambda \int_{-\infty}^{\infty} \frac{1}{\mu}\, d_\mu \int_0^1 f(y)\, \theta_n(x, y \mid \mu)\, dy \\ &\quad + \lambda^2 \int_{-\infty}^{\infty} \frac{1}{\mu(\mu - \lambda)}\, d_\mu \int_0^1 f(y)\theta_n(x, y \mid \mu)\, dy \\ &= f(x) + \lambda \int_0^1 K_n(x, y)\, f(y)\, dy \\ &\quad + \lambda^2 \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 \int_0^1 K_n(x, s)\theta_n(s, t \mid \mu)\, f(t)\, ds\, dt. \end{aligned} \tag{3.18}$$

[20]

---

[17] Cf. (C).

[18] This theorem is analogous to a result given by Carleman, for λ non-real, in (C; pp. 66–68).

[19] This is in accordance with (C; p. 26).

[20] (3.18) is analogous to a formula given in (C; p. 67) for the Type 1.A; (3.18) amounts to a rewriting in terms of Stieltjes integration of relations which are derived without difficulty from the classical theory.

Now

$$\lim_n \int_0^1 K_n(x, y)f(y)\, dy = \int_0^1 K(x, y)f(y)\, dy, \tag{3.18a}$$

since $f(y) \subset L_2$ and since

$$K_n(x, y) \to K(x, y), \qquad | K_n(x, y) | \leqq | K(x, y) | \subset L_2 \qquad \text{(in } y\text{)};$$

in fact, these conditions enable application of the theorem in connection with (3.5), (3.5a).

We write

$$C_n = \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 \int_0^1 \theta_n(s, t \mid \mu)\, K_n(x, s) f(t)\, ds\, dt = C'_n + C''_n, \tag{3.19}$$

where

$$C'_n = \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 \int_0^1 \theta_n(s, t \mid \mu)\, K(x, s) f(t)\, ds\, dt, \tag{3.19a}$$

$$C''_n = \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 \int_0^1 \theta_n(s, t \mid \mu)[K_n(x, s) - K(x, s)]\, f(t)\, ds\, dt. \tag{3.19b}$$

Let us express C'_n as

$$C'_n = C'_n(l) + R'_n(l), \tag{3.20}$$

where

$$C'_n(l) = \int_{-l}^{l} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 \int_0^1 \theta_n(s, t \mid \mu) K(x, s) f(t)\, ds\, dt, \tag{3.20a}$$

$$R'_n(l) = \left(\int_{-\infty}^{-l} + \int_{l}^{\infty} \right) \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 \int_0^1 \theta_n(s, t \mid \mu) K(x, s) f(t)\, ds\, dt. \tag{3.20b}$$

Keeping λ on the interval d and letting $l > | \lambda_0 | + h$, it is observed that, in $R'_n(l)$,

$$\left| \frac{1}{\mu - \lambda} \right| \leqq \frac{1}{l - |\lambda_0| - h};$$

consequently

$$| R'_n(l) |^2 \leqq \frac{1}{(l - | \lambda_0 | - h)^2} \int_0^1 | f(t) |^2\, dt \int_0^1 K^2(x, s)\, ds \leqq \epsilon^2 \qquad (\epsilon > 0) \tag{3.20c}$$

(for $l \geqq l(\epsilon, x)$), as can be deduced with the aid of a result in (C; p. 31).[21]

---

[21] This result gives an inequality for the variation of the double integral involved in (3.20b). In this connection, as well as on several occasions in the sequel it is to be noted that the formulas referred to in (C) are applicable not only for kernels of Type 1.A, but also for kernels of Type 1. B. This is true, of course, not for all results established in (C).

Consider now $C'_n(l)$ (3.20a),

$$C'_n(l) = A'_n(l) + B'_n(l), \tag{3.21}$$

where

$$A'_n(l) = \left(\int_{-l}^{\lambda_0-\omega} + \int_{\lambda_0+\omega}^{l}\right)\frac{1}{\mu-\lambda}\,d_\mu \int_0^1\int_0^1 \theta_n(s, t \mid \mu)K(x, s)f(t)\,ds\,dt, \tag{3.21a}$$

$$B'_n(l) = \int_\Delta \frac{1}{\mu-\lambda}\,d_\mu \int_0^1\int_0^1 \theta_n(s, t \mid \mu)K(x, s)f(t)\,ds\,dt. \tag{3.21b}$$

It is observed that in $A'_n(l)$

$$\left|\frac{1}{\mu-\lambda}\right| \leqq \frac{1}{\omega - h};$$

thus, for λ fixed on d the function of μ, $\dfrac{1}{\mu-\lambda}$, is continuous in μ for μ on each of the closed intervals

$$(-l, \lambda_0 - \omega), \qquad (\lambda_0 + \omega, l);$$

hence application of a result in (C)[22] will yield

$$\begin{aligned} \lim_n A'_n(l) &= A'(l) \\ &= \left(\int_{-l}^{\lambda_0-\omega} + \int_{\lambda_0+\omega}^{l}\right)\frac{1}{\mu-\lambda}\,d_\mu \int_0^1\int_0^1 \theta(s, t \mid \mu)\,K(x, s)\,f(t)\,ds\,dt. \end{aligned} \tag{3.21c}$$

Here

$$\theta(s, t \mid \mu) = \lim \theta_n(s, t \mid \mu) = \text{spectrum of } K(s, t);$$

it is understood that passage to the limit is through a suitable subsequence of values n.

In view of the expression for $\theta_n(s, t \mid \lambda)$ in terms of the characteristic functions, from (3.21b) one obtains

$$\begin{aligned} B'_n(l) &= \sum_\nu{}' \frac{1}{\lambda_{n,\nu} - \lambda}\int_0^1\int_0^1 K(x, s)\varphi_{r,\nu}(s)\varphi_{n,\nu}(t)f(t)\,ds\,dt \\ &= \sum_\nu{}' \frac{f_{n,\nu}}{\lambda_{n,\nu} - \lambda}\int_0^1 K(x, s)\varphi_{n,\nu}(s)\,ds \end{aligned} \tag{3.21d}$$

$[n = m_j; f_{n,j} = \text{Fourier coefficient of } f(s)]$.

In (3.21d) the sum is over those values ν for which the points $\lambda_{n,\nu}$ are in Δ. Now, by hypothesis, there are no points $\lambda_{n,\nu}$ in Δ; hence

$$B'_n(l) = 0. \tag{3.21e}$$

[22] (C; formula 45; p. 32).

In consequence of (3.20), (3.20c), (3.21) and (3.21e)

$$| C'_n - A'_n(l) | = | R'_n(l) | \leqq \epsilon \qquad (\text{for } l \geqq l(\epsilon, x)).$$

Hence, by (3.21c),

$$\lim_n C'_n = \left(\int_{-\infty}^{\lambda_0 - \omega} + \int_{\lambda_0 + \omega}^{\infty} \right) [\text{integrand from (3.21c)}]$$

$$(3.22) \qquad = \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu \int_0^1 \int_0^1 K(x, s)\theta(s, t \mid \mu) f(t)\, ds\, dt.$$

We turn our attention now to C''_n (3.19b),

$$(3.23) \qquad C''_n = A''_n + B''_n,$$

where

$$(3.23a) \quad B''_n = \int_\Delta \frac{1}{\mu - \lambda} d_\mu \int_0^1 \int_0^1 \theta_n(s, t \mid \mu)[K_n(x, s) - K(x, s)] f(t)\, ds\, dt,$$

$$(3.23b) \quad A''_n = \left(\int_{-\infty}^{\lambda_0 - \omega} + \int_{\lambda_0 + \omega}^{\infty} \right) [\text{integrand as in (3.23a)}].$$

It is observed that B''_n is $B'_n(l)$ (cf. (3.21b)) with $K(x, s)$ replaced by $K_n(x, s) - K(x, s)$; thus, in place of (3.21d) one obtains

$$B''_n = \sum_\nu{}' \frac{f_{n,\nu}}{\lambda_{n,\nu} - \lambda} \int_0^1 [K_n(x, s) - K(x, s)]\varphi_{n,\nu}(s)\, ds,$$

where the summation symbol has the same significance as in (3.21d). Hence

$$(3.23c) \qquad B''_n = 0.$$

With reference to (3.23b) it is noted that, inasmuch as λ is on d,

$$\left| \frac{1}{\mu - \lambda} \right| \leqq \frac{1}{\omega - h};$$

accordingly,

$$| A''_n | \leqq \frac{1}{\omega - h} \underset{(\lambda)}{\text{Var.}} \int_0^1 \int_0^1 \theta_n(s, t)[K_n(x, s) - K(x, s)] f(t)\, ds\, dt.$$

Whence, in view of (3.23) and (3.23c), an application of (C; p. 31) will yield

$$| C''_n |^2 = | A''_n |^2 \leqq \frac{1}{(\omega - h)^2} \int_0^1 | f(t) |^2\, dt \int_0^1 [K_n(x, s) - K(x, s)]^2\, ds,$$

which implies that

$$(3.24) \qquad \lim_n C''_n = 0.$$

By virtue of (3.19), (3.24) and (3.22) it can now be asserted that

$$(3.25) \qquad \lim_i C_{m_i} = \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu \int_0^1 \int_0^1 K(x, s)\theta(s, t \mid \mu) f(t)\, ds\, dt = C$$

for a suitable subsequence (m_j) and for λ on d. Thus, on taking account of (3.18a) and of (3.18), it is inferred that for almost all x (on 0, 1)), $K(x, y)$ being of Type 1.B, we have

$$\begin{aligned}(3.26)\qquad \lim_j \varphi_{m_j}(x) = \varphi(x) &= f(x) + \lambda \int_0^1 K(x, y) f(y)\, dy \\ &+ \lambda^2 \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 \int_0^1 K(x, s)\theta(s, t \mid \mu) f(t)\, ds\, dt\end{aligned}$$

$$[\lambda \text{ on } d;\ \theta(s, t \mid \mu) = \text{spectrum of } K(s, t) \text{ corresponding to } (m_j)].$$

The function $\varphi(x)$ so defined will $\subset L_2$ and will be a solution of (1.1) for almost all x on $(0, 1)$.

With λ on d, it is possible to show that (3.26) implies (3.17). A similar result is established, following the procedure given above, when $K(x, y)$ is of Type 1.A. Thus, *the theorem is seen to hold as stated.*

Suppose now that $K(x, y)$ is of Type 1.C, L denoting a corresponding linear operator. With $\theta_n(x, y \mid \lambda)$ defined as stated subsequent to Theorem 3.3, form the function

$$(3.27)\qquad \int_0^x \int_0^y \theta_n(x, y \mid \lambda)\, dx\, dy = \Omega_n(x, y \mid \lambda).$$

A subsequence m_j of the sequence (2.2) (involved in the definition of G) can be found so that the limit

$$(3.27a)\qquad \lim_j \Omega_{m_j}(x, y \mid \lambda) = \Omega(x, y \mid \lambda)$$

exists and possesses most of the essential properties enjoyed by the function so denoted in (C)—in Carleman's treatment of kernels of the kind somewhat less general than Type 1.C.[23] Following the procedure given in (C) (Chapter IV) *it can be shown (for the kernels of Type* 1.C) *that, corresponding to every limiting function* $\Omega(x, y \mid \lambda)$ (3.27a), *the equation*

$$(3.28)\qquad L_x(\xi \mid \varphi(x)) - \lambda \int_0^1 L_x(\xi \mid K(x, y))\varphi(y)\, dy = L_x(\xi \mid f(x))$$

will possess a solution $\subset L_2$ *of the form*

$$(3.28a)\qquad \varphi(x) = f(x) + \lambda \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 f(y) \frac{\partial}{\partial y} \Omega(x, y \mid \mu)\, dy$$

(*almost everywhere on* (0, 1)), *provided* λ *is non-real.*

An analogue to Theorem 3.3 will be as follows.

THEOREM 3.4. *With* $K(x, y)$ *of Type* 1.C *and* λ *on an interval* d, *a solution* $\varphi(x) \subset L_2$ *of the equation* (3.28) *is given by the formula* (3.28a), *where* $\Omega(x, y \mid \mu)$ *is a function of the form* (3.27a).

---

[23] Cf. Chapter IV of (C). Also see Trjitzinsky, loc. cit.

To establish this result we first note that in consequence of the classical formula for solution $\varphi_{m_j}(x)$, $\subset L_2$, of (1.10) (with $n = m_j$) we have

$$\varphi_{m_j}(x) = f(x) + \lambda \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu \int_0^1 f(y) \frac{\partial}{\partial y} \Omega_{m_j}(x, y \mid \mu)\, dy, \tag{3.29}$$

where $\Omega_{m_j}(x, y \mid \mu)$ is defined by (3.27). In view of the preceding, the m_j, involved in (3.27a), may be selected so that

$$\int_0^1 | \varphi_{m_j}(x) |^2\, dx \leqq q^2, \qquad \varphi_{m_j}(x) \rightarrow \varphi(x) \qquad \text{(weakly)}. \tag{3.29a}$$

We also shall have

$$L_x(\xi \mid \varphi_{m_j}(x)) - \lambda \int_0^1 L_x(\xi \mid K_{m_j}(x, y))\varphi_{m_j}(y)\, dy = L_x(\xi \mid f(x)). \tag{3.29b}$$

The function $\varphi(x)$, referred to in (3.29a), will be a solution of (3.28). Accordingly, it is observed that the theorem will be established as soon as it is shown that

$$\lim_j \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu h_j(x, \mu) = \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda}\, d_\mu h(x, \mu), \tag{3.30}$$

where

$$\begin{aligned} h_j(x, \mu) &= \int_0^1 f(y) \frac{\partial}{\partial y} \Omega_{m_j}(x, y \mid \mu)\, dy, \\ h(x, \mu) &= \int_0^1 f(y) \frac{\partial}{\partial y} \Omega(x, y \mid \mu)\, dy. \end{aligned} \tag{3.30a}$$

Now, for μ on Δ

$$\theta_{m_j}(x, y \mid \mu) = 0.$$

Hence

$$\Omega_{m_j}(x, y \mid \mu) = 0, \qquad \Omega(x, y \mid \mu) = 0 \qquad (\mu \text{ on } \Delta).$$

Whence (3.30) will hold if

$$\lim_j \int_{I(\Delta)} \frac{1}{\mu - \lambda}\, d_\mu(h(x, \mu) - h_j(x, \mu)) = 0, \tag{3.30b}$$

where

$$I(\Delta) = (-\infty, +\infty) - \Delta. \tag{3.30c}$$

Write

$$\int_{I(\Delta)} \frac{1}{\mu - \lambda}\, d_\mu(h(x, \mu) - h_j(x, \mu)) = \int_{I(l \mid \Delta)} \cdots + \left(\int_{-\infty}^{-l} + \int_l^{\infty} \right) \cdots \tag{3.30d}$$

with

$$I(l \mid \Delta) = (-l, l) - \Delta,$$

where l is sufficiently great so that the interval $(-l, l)$ contains Δ.

For μ on $I(l \mid \Delta)$

$$\left|\frac{1}{\mu - \lambda}\right| \leqq \frac{1}{\omega - h} \qquad (\lambda \text{ on } d);$$

hence $1/(\mu - \lambda)$ is continuous in μ for μ on $I(l \mid \Delta)$. Thus by a theorem of *Helly*

$$\lim_j \int_{I(l \mid \Delta)} \frac{1}{\mu - \lambda} d_\mu(h(x, \mu) - h_j(x, \mu)) = 0, \tag{3.31}$$

inasmuch as $h_j(x, \mu) \to h(x, \mu)$ and

$$\text{Var. } h_j(x, \mu) \leqq \alpha \qquad (\alpha \text{ independent of } j);$$

the latter inequality being essentially a consequence of some of Carleman's developments. On the other hand,

$$\begin{aligned} &\left(\int_{-\infty}^{-l} + \int_{l}^{\infty}\right) \frac{1}{\mu - \lambda} d_\mu(h(x, \mu) - h_j(x, \mu)) \\ &\qquad \leqq \frac{1}{d(\lambda, l)} \text{Var. } (h(x, \mu) - h_j(x, \mu)) \leqq \frac{2\alpha}{d(\lambda, l)}, \end{aligned} \tag{3.32}$$

where

$$d(\lambda, l) = \text{lesser of } | l - \lambda |, \qquad | l + \lambda |.$$

We keep x (on (0, 1)) and λ (on d) fixed. Choose l sufficiently great so that

$$\left|\left(\int_{-\infty}^{-l} + \int_{l}^{\infty}\right) \cdots \right| \leqq \frac{\epsilon}{2},$$

which is possible by virtue of (3.32). Making use of (3.31) we then take $j(\epsilon)$ so that

$$\left|\int_{I(l \mid \Delta)} \cdots \right| \leqq \frac{\epsilon}{2} \qquad (\text{for all } j \geqq j(\epsilon)).$$

Thus, by (3.30d)

$$\left|\int_{I(\Delta)} \frac{1}{\mu - \lambda} d_\mu(h(x, \mu) - h_j(x, \mu))\right| \leqq \epsilon \qquad (j \geqq j(\epsilon)).$$

Accordingly, it is observed that (3.30b) holds. The truth of Theorem 3.4 is now evident.

4. Questions of uniqueness

Results relating to uniqueness properties for kernels of Type 1.A, for λ non-real, have been given by Carleman.[24] We shall establish the corresponding results, for kernels of Types 1.A, 1.B, when λ is in S_d, with S_d specified by Definition 2.1.

---

[24] (C; Chapter II).

We shall first prove the following theorem.

THEOREM 4.1. *Let $K(x, y)$ be a kernel of Type 1.B. Let λ be in S_d. With (n_j) designating a sequence (2.2), involved in the definition of the interval d, there exists a subsequence (m_j) independent of $f(x)$ so that, designating by $\varphi_{m_j}(x)$ the solution of*

$$\varphi_{m_j}(x) = \lambda \int_0^1 K_{m_j}(x, y)\varphi_{m_j}(y)\, dy + f(x) \qquad (f(x) \subset L_2), \tag{4.1}$$

we have

$$\lim_j \varphi_{m_j}(x) = \varphi(x) = T_\lambda(f \mid x) \tag{4.1a}$$

for all $f(x) \subset L_2$; moreover, $\varphi(x)$ will satisfy (1.1). $T_\lambda(f \mid x)$ will be a linear functional of $f(x)$ such that

$$\int_0^1 f_1(x) T_\lambda(f_2 \mid x)\, dx = \int_0^1 f_2(x) T_\lambda(f_1 \mid x)\, dx \qquad (\lambda \text{ in } S_d), \tag{4.2}$$

whenever $f_1(x), f_2(x) \subset L_2$.

For kernels of Type 1.A a similar result will hold for some subsequence (δ_i') (cf. (2.4)), with $\varphi_{m_j}(x)$ and $K_{m_j}(x, y)$ replaced by $\varphi^{(\delta_i')}(x)$ and $K^{(\delta_i')}(x, y)$, respectively.

To demonstrate the above we confine ourselves to the Type 1.B and denote by ${}_i\varphi_{n_j}$ the solutions $\subset L_2$ of

$${}_i\varphi_{n_j}(x) = \lambda \int_0^1 K_{n_j}(x, y)\, {}_i\varphi_{n_j}(y)\, dy + f_i(x) \qquad [f_i(x) \subset L_2;\ i = 1, 2]. \tag{4.3}$$

Then, as is known,

$$\int_0^1 {}_2\varphi_{n_j}(x) f_1(x)\, dx = \int_0^1 {}_2\varphi_{n_j}(x) f_2(x)\, dx. \tag{4.3a}$$

We have, with λ in S_d,

$$\int_0^1 |{}_1\varphi_{n_j}(x)|^2\, dx \leqq M_1(\lambda), \qquad \int_0^1 |{}_2\varphi_{n_j}(x)|^2\, dx \leqq M_2(\lambda). \tag{4.4}$$

In consequence of Lemma 2.1, for λ on d we may let

$$M_i(\lambda) = \left(1 + \frac{\lambda_1}{\omega - h}\right)^2 \int_0^1 |f_i(x)|^2\, dx = q_i^2 \qquad (i = 1, 2) \tag{4.4a}$$

$$[\lambda_1 = \text{upper bound of } |\lambda| \text{ for } \lambda \text{ on } d].$$

On the other hand, in view of (C), *for λ non-real* one may write

$$M_i(\lambda) = \frac{|\lambda|^2}{\beta^2} \int_0^1 |f_i(x)|^2\, dx \qquad (i = 1, 2), \tag{4.4b}$$

where β is the imaginary part of λ. Thus $M_i(\lambda)$ is finite for every λ in S_d.

We have shown previously how one may select a subsequence (m_j) of (n_j) so that

$$\lim_j \varphi_{m_j}(x) = \varphi(x)$$

exists almost everywhere, $\subset L_2$ and constitutes a solution of (1.1). The subsequence (m_j) will be selected independent of f.[25] Hence the relation (4.1a) may be now asserted. It remains to establish (4.2). For the latter purpose we apply to (4.3a), where n_j is replaced by m_j, the result in connection with (3.5), (3.5a). This application is possible inasmuch as for λ in S_d we have the inequalities (4.4), where the second members are finite (for every λ fixed in S_d). Hence passing to the limit and taking account of the already established formula (4.1a) one obtains (4.2) *for every* λ *in* S_d.

Introducing certain slight modifications in the above developments, similar results are established for kernels of Type 1.A.

From Theorem 4.1 it follows that if, for λ fixed in S_d, the equation (1.1) has only one solution $\varphi(x) \subset L_2$, then $\varphi(x) = T_\lambda(f \mid x)$ and (4.2) will hold for all $f_1(x), f_2(x) \subset L_2$.

THEOREM 4.2. *Let $K(x, y)$ be of Type* 1.A *or of Type* 1.B. *If for $\lambda = \lambda_1$ fixed in S_d the only solution $\subset L_2$ of the homogeneous equation*

$$\varphi(x) = \lambda \int_0^1 K(x, y)\varphi(y)\, dy \tag{4.5}$$

is zero,[26] *the same will be true for all non-real values of* λ.

In view of Theorem II_2 (C; p. 55) it is known that if λ_1 is non-real the above will certainly hold. Thus take λ_1 on d. Suppose now that the Theorem is not true. There will then exist a non-real value λ and a corresponding function $\varphi(x)$ for which

$$\int_0^1 |\varphi(x)|^2\, dx \neq 0, \qquad \varphi(x) = \lambda \int_0^1 K(x, y)\varphi(y)\, dy. \tag{4.6}$$

For this function we have

$$\varphi(x) - \lambda_1 \int_0^1 K(x, y)\varphi(y)\, dy = \left(1 - \frac{\lambda_1}{\lambda}\right)\varphi(x). \tag{4.6a}$$

The equation

$$\varphi(x) - \lambda_1 \int_0^1 K(x, y)\varphi(y)\, dy = f_1(x) \qquad (f_1(x) \subset L_2) \tag{4.6b}$$

cannot have two distinct solutions since otherwise their difference $w(x)$ will satisfy

[25] The possibility of such a choice follows from a remark in the foot-note of (C; p. 56).

[26] Almost everywhere, when $K(x, y)$ is of Type 1.B; for $x \neq \xi_\nu(\nu = 1, 2, \cdots)$, when $K(x, y)$ is of Type 1.A.

$$w(x) - \lambda_1 \int_0^1 K(x, y)w(y)\,dy = 0, \qquad \int_0^1 |w(x)|^2\,dx \neq 0, \tag{4.6c}$$

contrary to hypothesis. Thus (4.6b) *has a unique solution* $\subset L_2$; hence in view of (4.6a) and of Theorem 4.1 the function $\varphi(x)$, involved in (4.6), satisfies the functional relation

$$\varphi(x) = T_{\lambda_1}\left(\left(1 - \frac{\lambda_1}{\lambda}\right)\varphi \,\middle|\, x\right). \tag{4.7}$$

On taking the conjugates of the members of (4.6a) it is deduced that

$$\bar{\varphi}(x) - \lambda_1 \int_0^1 K(x, y)\bar{\varphi}(y)\,dy = \left(1 - \frac{\lambda_1}{\lambda}\right)\bar{\varphi}(x), \tag{4.8}$$

so that in consequence of Theorem 4.1 and of the italics subsequent to (4.6c)

$$\bar{\varphi}(x) = T_{\lambda_1}\left(\left(1 - \frac{\lambda_1}{\lambda}\right)\bar{\varphi} \,\middle|\, x\right). \tag{4.8a}$$

Application of the same Theorem to (4.6a), (4.8) will yield

$$\int_0^1 \left(1 - \frac{\lambda_1}{\lambda}\right)\varphi(x)T_{\lambda_1}\left(\left(1 - \frac{\lambda_1}{\lambda}\right)\bar{\varphi} \,\middle|\, x\right)dx = \int_0^1 \left(1 - \frac{\lambda_1}{\lambda}\right)\bar{\varphi}(x)T_{\lambda_1}\left(\left(1 - \frac{\lambda_1}{\lambda}\right)\varphi \,\middle|\, x\right)dx.$$

Thus, by (4.7) and (4.8a)

$$\left(1 - \frac{\lambda_1}{\lambda}\right)\int_0^1 |\varphi(x)|^2\,dx = \left(1 - \frac{\lambda_1}{\lambda}\right)\int_0^1 |\varphi(x)|^2\,dx. \tag{4.9}$$

By virtue of (4.6) from (4.9) it is inferred that λ must be real, which is contrary to hypothesis. The theorem is accordingly established.

In (C) *Carleman* gave conditions under which kernels are of Class I; that is, possess the property that for some (and hence all) non-real values of λ the homogeneous equation has zero as the only solution $\subset L_2$. In view of Theorem 4.2 the following may be asserted.

In order that a kernel $K(x, y)$ (of Type 1.A *or of Type* 1.B*) be of Class* I *it is sufficient that for some value of λ on d the only solution ($\subset L_2$) of the homogeneous equation should be zero.*

We shall now prove the following theorem.

THEOREM 4.3. *Let $K(x, y)$ be of Type* 1.A *or of Type* 1.B. *The number m of distinct*[27] *solutions ($\subset L_2$) of*

$$\varphi(x) = \lambda \int_0^1 K(x, y)\varphi(y)\,dy \tag{4.10}$$

[27] 'Distinct', here and in the sequel signifies 'linearly independent.'

for any λ fixed on d, is equal to or is greater than the number n of distinct solutions for non-real values of λ.

We first note that the number of distinct solutions is the same for all non-real values of λ.[28] Suppose the theorem is not true. We then have n, $n > m$, distinct solutions $\varphi_\nu^*(x)$,

$$\varphi_\nu^*(x) = \lambda^* \int_0^1 K(x, y)\varphi_\nu^*(y)\,dy \qquad (\nu = 1, \cdots n), \tag{4.11}$$

for a non-real value λ^*. On the other hand, for a fixed λ on d, there exist by hypothesis m distinct solutions $\varphi_j(x)$,

$$\varphi_j(x) = \lambda \int_0^1 K(x, y)\varphi_j(y)\,dy \qquad (j = 1, \cdots m). \tag{4.11a}$$

Inasmuch as λ and $K(x, y)$ are real we may suppose that the $\varphi_j(x)$ are real. Form a solution $\varphi^*(x)$ for λ^*,

$$\varphi^*(x) = \sum_{\nu=1}^{n} c_\nu \varphi_\nu^*(x) \qquad (\sum_\nu |c_\nu|^2 \neq 0); \tag{4.12}$$

then

$$\int_0^1 |\varphi^*(x)|^2\,dx \neq 0. \tag{4.12a}$$

In addition, let the c_ν be so chosen that

$$\int_0^1 \varphi^*(x)\bar{\varphi}(x)\,dx = 0 \tag{4.12b}$$

for all solutions $\varphi(x)$ satisfying (4.10) *for the value* λ. The possibility of securing (4.12b) is inferred from the following considerations. One has

$$\varphi(x) = \sum_{j=1}^{n} d_j \varphi_j(x).$$

Thus, (4.12b) will be satisfied if

$$\int_0^1 \varphi^*(x)\bar{\varphi}_j(x)\,dx = 0 \qquad (j = 1, \cdots m).$$

That is, we need to choose the c_ν (not all zero) so that

$$\sum_{\nu=1}^{n} c_\nu \int_0^1 \varphi_\nu^*(x)\bar{\varphi}_j(x)\,dx = 0 \qquad (j = 1, 2, \cdots m).$$

Such a choice is possible since $m < n$.

For the function (4.12) we have

$$\varphi^*(x) - \lambda^* \int_0^1 K(x, y)\varphi(y)\,dy = 0, \quad \bar{\varphi}^*(x) - \bar{\lambda}^* \int_0^1 K(x, y)\bar{\varphi}^*(y)\,dy = 0; \tag{4.13}$$

[28] (C; p. 58).

these relations may be written in the form

$$\text{(4.13a)}\qquad \begin{aligned} \varphi^*(x) - \lambda \int_0^1 K(x, y)\varphi^*(y)\,dy &= \left(1 - \frac{\lambda}{\lambda^*}\right)\varphi^*(x) = f(x), \\ \bar{\varphi}^*(x) - \lambda \int_0^1 K(x, y)\bar{\varphi}^*(y)\,dy &= \left(1 - \frac{\lambda}{\bar{\lambda}^*}\right)\bar{\varphi}^*(x) = \bar{f}(x). \end{aligned}$$

In view of Theorem 4.1 it is accordingly deduced that

$$\text{(4.14)}\qquad \begin{aligned} \varphi^*(x) &= T_\lambda(f \mid x) + \psi_1(x), \\ \bar{\varphi}^*(x) &= T_\lambda(\bar{f} \mid x) + \psi_2(x), \end{aligned}$$

where $\psi_1(x)$, $\psi_2(x)$ are certain solutions of (4.10) for λ. The operator $T_\lambda(h \mid x)$ (λ on d) may be so determined that

$$\overline{T}_\lambda(h \mid x) = T_\lambda(\bar{h} \mid x) \qquad (\text{all } h(x) \subset L_2);$$

we then have

$$\text{(4.14a)}\qquad \psi_2(x) = \bar{\psi}_1(x).$$

A further application of Theorem 4.1 will yield

$$\int_0^1 f(x)T_\lambda(\bar{f} \mid x)\,dx = \int_0^1 \bar{f}(x)T_\lambda(f \mid x)\,dx;$$

that is, in consequence of (4.14) and (4.14a),

$$\left(1 - \frac{\lambda}{\lambda^*}\right)\int_0^1 \varphi^*(x)[\bar{\varphi}^*(x) - \bar{\psi}_1(x)]\,dx = \left(1 - \frac{\lambda}{\bar{\lambda}^*}\right)\int_0^1 \bar{\varphi}^*(x)[\varphi^*(x) - \psi_1(x)]\,dx.$$

By virtue of the statement with respect to (4.12b)

$$\int_0^1 \varphi^*(x)\bar{\psi}_1(x)\,dx = 0, \qquad \int_0^1 \bar{\varphi}^*(x)\psi_1(x)\,dx = 0.$$

Whence

$$\left(1 - \frac{\lambda}{\lambda^*}\right)\int_0^1 |\varphi^*(x)|^2\,dx = \left(1 - \frac{\lambda}{\bar{\lambda}^*}\right)\int_0^1 |\varphi^*(x)|^2\,dx,$$

which by (4.12a) implies that $\lambda^* = \bar{\lambda}^*$. Thus, λ^* must be real, contrary to the hypothesis made in connection with (4.11).. Whence the theorem holds as stated.

Theorem 4.3 will still hold when the number of distinct solutions is infinite.

If $K(x, y)$ is of Type 1.C and if $L_x(\xi \mid h(x))$ is a linear operator associated with $K(x, y)$, in accordance with the conditions (1.9), $\cdots$ (1.9d), theorems analogous to Theorems 4.1, 4.2, 4.3 will hold for the equation

$$L_x(\xi \mid \varphi(x)) - \lambda \int_0^1 L_x(\xi \mid K(x, y))\varphi(y)\,dy = L_x(\xi \mid f(x)),$$

provided $L_x(\xi \mid h(x))$ is real for $h(x), \subset L_2$, real and provided

$$\bar{L}_x(\xi \mid h(x)) = L_x(\xi \mid \bar{h}(x)).$$

5. A further extension and examples

Under certain circumstances solutions of the non-homogeneous equation may exist for λ real even if λ does not belong to any interval d of the description given in Definition 2.1. In this connection the following Lemma will be helpful.

LEMMA 5.1. *Let $K(x, y)$ be of Type 1.B. Designate by G a set*

$$G = \{\lambda_{n,\nu}\} \qquad (n = n_j\,;\, j, \nu = 1, 2, \cdots), \tag{5.1}$$

where (n_j) forms an infinite subsequence of (n). Let T be a set on a finite part of the real axis in the λ-plane, no point of T being coincident with a point of G.[29] *With the equation (1.1) in view, write*

$$f_{n,\nu} = \int_0^1 f(x)\varphi_{n,\nu}(x)\,dx. \tag{5.2}$$

Form the series

$$s_n^2(\lambda) = \sum_{\nu=1}^{\infty} \left| \frac{f_{n,\nu}}{\lambda_{n,\nu} - \lambda} \right|^2 \qquad (n = n_j; j = 1, 2, \ldots). \tag{5.3}$$ [30]

If

$$| s_n(\lambda) | \leqq s(\lambda) \qquad (s(\lambda) \textit{ independent of } n = n_j\,;\, n_1 < n_2 < \cdots) \tag{5.3a}$$

where $s(\lambda)$ is finite for every λ in T, then

$$\int_0^1 | \varphi_n(x) |^2\,dx \leqq m^2(\lambda) \qquad (\lambda \text{ in } T; n = n_1, n_2, \cdots), \tag{5.4}$$

with $\varphi_n(x)$ constituting a solution $\subset L_2$ of (2.8) and

$$m(\lambda) = | \lambda | s(\lambda) + \left[\int_0^1 | f(x) |^2\,dx \right]^{\frac{1}{2}}. \tag{5.4a}$$

To establish this Lemma we note that by (2.8a)

$$\int_0^1 | \varphi_n(x) |^2\,dx = \lambda w_n + \int_0^1 f(x)\bar{\varphi}_n(x)\,dx, \tag{5.5}$$

where

$$w_n = \int_0^1 \int_0^1 K_n(x, y)\varphi_n(y)\bar{\varphi}_n(x)\,dx\,dy.$$

[29] It may occur that every point of T is a limiting point of G.

[30] For n fixed the $\lambda_{n,\nu}$ have no limiting points on the finite part of the axis of reals; λ (in T) is distinct from $\lambda_{n,\nu}(\nu = 1, 2, \cdots)$ and is not a limiting point of the $\lambda_{n,\nu}$; hence λ is at a positive distance from the set $(\lambda_{n,\nu})(\nu = 1, 2, \cdots ; n \text{ fixed})$. Clearly, every series (5.3) will converge for λ in T, inasmuch as $| f_{n,1} |^2 + | f_{n,2} |^2 + \cdots$ converges.

In consequence of the first inequality (2.11) and of (5.3)

$$|w_n|^2 \leqq s_n^2(\lambda) \int_0^1 |\varphi_n(x)|^2 dx \qquad (\lambda \text{ in } T). \tag{5.5a}$$

Whence by virtue of (5.5)

$$\int_0^1 |\varphi_n(x)|^2 dx \leqq |\lambda| |w_n| + \left[\int_0^1 |f(x)|^2 dx\right]^{\frac{1}{2}} \left[\int_0^1 |\varphi_n(x)|^2 dx\right]^{\frac{1}{2}}$$
$$\leqq \left\{|\lambda| s_n(\lambda) + \left[\int_0^1 |f(x)|^2 dx\right]^{\frac{1}{2}}\right\} \left[\int_0^1 |\varphi_n(x)|^2 dx\right]^{\frac{1}{2}},$$

which yields

$$\int_0^1 |\varphi_n(x)|^2 dx \leqq \left\{|\lambda| s_n(\lambda) + \left[\int_0^1 |f(x)|^2 dx\right]^{\frac{1}{2}}\right\}^2 \tag{5.6}$$

for λ in T and $n = n_1, n_2, \cdots$. Finally, (5.4) will follow from (5.6) in consequence of the hypothesis (5.3a). This demonstrates the Lemma.

With the aid of the above Lemma the following can be proved.

THEOREM 5.1. *The results stated in the previous sections for kernels of Type* 1.B *will hold (with obvious modifications) when intervals d (cf. Definition* 2.1) *are replaced by sets T (T defined in Lemma* 5.1) *for which* (5.3a) *holds.*

In order to see the truth of this statement one needs only to repeat the developments previously given, employing whenever necessary the inequality (5.4). Analogous facts will hold for kernels of Types 1.A and 1.B.

In applications equations often occur in the form

$$\varphi(x) = \int_0^1 K(x, y)\varphi(y)\, dy + f(x),$$

where $K(x, y)$ is a kernel of one of the Types studied above. Our results could be applied whenever an interval d (Definition 2.1) or a set T, as described in Lemma 5.1, can be found so that $\lambda = 1$ belongs to d or to T, as the case may be.

Under suitable conditions, many of the above results will hold when $K(x, y)$ is approximated by regular (that is, belonging to L_2 in (x, y)) kernels, even when the modes of approximation are different from those previously used.

When the approximating kernels can be taken of positive type[31] then the characteristic values are all positive and, clearly, any interval

$$(c_1 \leqq \lambda \leqq c_2) \qquad (c_1 < c_2 < 0)$$

could be used as d.

A class of singular kernels may be constructed as follows. Let

$$\psi_{m,\nu}(x) \qquad (m, \nu = 1, 2, \cdots) \tag{5.7}$$

[31] A kernel $A(x, y)$ is of positive type if $\iint A(x, y)\varphi(x)\varphi(y)\, dx\, dy \geqq 0$ for all $\varphi(x) \subset L_2$.

be real functions $\subset L_2$ for x on $(0, 1)$ and such that

$$\int_0^1 \psi_{m,\nu}(x)\psi_{m,j}(x)\,dx = 0 \qquad (\nu \neq j), \tag{5.7a}$$

$$\int_0^1 \psi_{m,\nu}^2(x)\,dx = 1 \qquad (m, \nu = 1, 2, \ldots), \tag{5.7b}$$

$$\int_0^1 \psi_{m,\nu}(x)\psi_{p,j}(x)\,dx = 0 \quad (m \neq p;\ \nu, j = 1, 2, \ldots).[32] \tag{5.7c}$$

Suppose real numbers $\lambda_{m,\nu}$ are such that the series

$$S_m = \sum_\nu \frac{1}{\lambda_{m,\nu}^2}, \qquad \sum_m \sum_\nu \frac{1}{\lambda_{m,\nu}^2}\psi_{m,\nu}^2(x), \qquad \sum_\nu \frac{\psi_{m,\nu}^2(x)}{|\lambda_{m,\nu}|} \tag{5.8}$$

converge (for almost all x on $(0, 1)$), while the series

$$S = \sum_{m=1}^{\infty} S_m \tag{5.8a}$$

diverges.

Form the functions $g_m(x, y)$,

$$g_m(x, y) = \sum_\nu \frac{1}{\lambda_{m,\nu}} \psi_{m,\nu}(x)\psi_{m,\nu}(y). \tag{5.9}$$

The integrals

$$\int_0^1\int_0^1 g_m^2(x, y)\,dx\,dy = S_m \qquad (m = 1, 2, \cdots) \tag{5.9a}$$

will exist. The $g_m(x, y)$ will be regular kernels; moreover, they will form an orthogonal set. The kernels

$$K_{(n)}(x, y) = g_1(x, y) + \cdots + g_n(x, y) \qquad (n = 1, 2, \cdots) \tag{5.9b}$$

will also be regular; in fact, the integrals

$$\int_0^1\int_0^1 K_{(n)}^2(x, y)\,dx\,dy = S_1 + \cdots + S_n \tag{5.9c}$$

exist. By (5.7a), (5.7b)

$$\int_0^1 g_m^2(x, y)\,dy = \sum_\nu \frac{1}{\lambda_{m,\nu}^2}\psi_{m,\nu}^2(x).$$

Thus

$$\int_0^1 K_{(n)}^2(x, y)\,dy = \sum_{m=1}^{n}\int_0^1 g_m^2(x, y)\,dy = \sum_{m=1}^{n}\sum_{\nu=1}^{\infty}\frac{1}{\lambda_{m,\nu}^2}\psi_{m,\nu}^2(x). \tag{5.9d}$$

---

[32] Examples of such functions are easily given.

The kernel

$$K(x, y) = \sum_{m=1}^{\infty} g_m(x, y) = \lim_{\nu} K_{(n)}(x, y) \tag{5.10}$$

will be singular of Type 1.B. The characteristic values for the approximating kernel $K_n(x, y)$ will be the numbers

$$\lambda_{m,\nu} \qquad (m = 1, \cdots n; \nu = 1, 2, \cdots);$$

accordingly, these values will be at our disposal, subject to the general conditions stated above.

We shall obtain a kernel which is merely measurable by dropping the requirement that the double series in (5.8) should converge.

6. Equations of the first kind

The study of integral equations of the first kind

$$\int_0^1 K(x, y)\varphi(y)\, dy = f(x) \qquad [\text{real } f(x) \subset L_2;\, x \text{ on } (0, 1)], \tag{6.1}$$

where $K(x, y)$ is of Type 1.A or 1.B, is reducible to that of another integral equation (also of the first kind but with a kernel which is not necessarily symmetric), to which the classical methods of *Picard* and *Lauricella* are applicable.[33] On the other hand, when $K(x, y)$ is of Type 1.C the integral

$$\int_0^x K(t, y)\, dt$$

does not need to exist; consequently the reduction, referred to above, will be in general impossible for kernels of Type 1.C. Accordingly, *in the sequel it will be supposed that $K(x, y)$ is of Type* 1.C; moreover, in place of (6.1) we shall study the related functional equation

$$\int_0^1 L_x(\xi \mid K(x, y))\varphi(y)\, dy = L_x(\xi \mid f(x)) \qquad (f(x) \subset L_2), \tag{6.2}$$

where $L_x(\xi \mid h(x))$ is an operator [associated] with $K(x, y)$, subject to the conditions (1.9), $\cdots$ (1.9d). It is conceivable that in some cases

$$L_x(\xi \mid K(x, y)) \subset L_1 \quad (\text{in } \xi), \qquad L_x(\xi \mid f(x)) \subset L_1 \quad (\text{in } \xi); \tag{6.3}$$

this would enable reduction of (6.2) to an equation to which the methods of Picard and Lauricella are applicable. However, (6.3) will hold not always. Accordingly, development of a different procedure for the study of (6.2) is desirable.

---

[33] References to the papers of these authors, as well as certain other facts concerning equations of the first kind are found in J. Soula, *L'équation intégrale de première espèce à limites fixes et les fonctions permutables à limites fixes* [Mémorial des Sciences Math., 80 (1936)]. For the device referred to in the text see pp. 3–4.

Let $K_n(x, y)$ be the kernel defined in (1.4). This kernel will be regular and will possess characteristic values $\lambda_{n,\nu}$ and characteristic functions $\varphi_{n,\nu}(x)$. In terms of these numbers and functions we form the spectrum $\theta_n(x, y \mid \lambda)$ (cf. the text subsequent to (3.17)) of $K_n(x, y)$. While for kernels of Type 1.B it is possible to select a subsequence $(\theta_{n_j}(x, y \mid \lambda))$ converging to a limiting function $\theta(x, y \mid \lambda)$, termed spectrum of $K(x, y)$, for kernels of Type 1.C existence of a spectrum cannot be asserted.[34] On writing, as before,

$$\Omega_n(x, y \mid \lambda) = \int_0^x \int_0^y \theta_n(x, y \mid \lambda)\, dx\, dy \tag{6.4}$$

and on recalling that there exists a subsequence (n_j) of (n) for which the limit

$$\lim_j \Omega_{n_j}(x, y \mid \lambda) = \Omega(x, y \mid \lambda) \tag{6.4a}$$

exists for almost all x, y in the square

$$0 \leqq x, y \leqq 1$$

and for all real values λ, except perhaps for $\lambda = \rho_1, \rho_2, \cdots$ (for which values Ω may be discontinuous in λ), it is also to be noted that

$$\underset{(\lambda)}{\text{Var.}}\, \Omega(x, y \mid \lambda) \leqq (xy)^{\frac{1}{2}}, \qquad \Omega(x, y \mid 0) = 0; \tag{6.4b}$$

$$\mid \Omega(x', y' \mid \lambda) - \Omega(x, y \mid \lambda) \mid \leqq [\mid y' - y \mid]^{\frac{1}{2}} + [\mid x' - x \mid]^{\frac{1}{2}}. \tag{6.4c}$$

In accordance with the terminology of *Carleman* it will be said that Ω is '*closed*', if the generalized Bessel's inequality (which holds for kernels of Type 1.C),

$$\int_{-\infty}^{\infty} d_\lambda \int_0^1 h(x) \left(\frac{\partial}{\partial x} \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda)\, dy \right) dx \leqq \int_0^1 h^2(x)\, dx, \tag{6.5}$$

is valid with the equality sign (for all $h(x) \subset L_2$); one then has

$$h(x) = \frac{d}{dx} \int_{-\infty}^{\infty} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda)\, dy \tag{6.5a}$$

for almost all x on (0, 1).

We shall establish the following theorem.

THEOREM 6.1. *Let $K(x, y)$ be of Type 1.C, with $K_n(x, y)$ designating the kernel (1.4), the characteristic numbers and functions of $K_n(x, y)$ being the $\lambda_{n,\nu}(\nu = 1, 2, \cdots)$ and the $\varphi_{n,\nu}(x)$ $(\nu = 1, 2, \cdots)$, respectively.*

Suppose that for a subsequence (n_j) of (n) each sequence

$$\varphi_{n_j,1}(x), \varphi_{n_j,2}(x), \cdots \tag{6.6}$$

forms a complete set and that the series

$$\Gamma_{n_j} = \sum_\nu \lambda_{n_j,\nu}^2 f_{n_j,\nu}^2 \qquad \left[f_{n,\nu} = \int_0^1 f(x)\varphi_{n,\nu}(x)\, dx \right] \tag{6.6a}$$

[34] Cf. (C), where kernels less general than those of Type 1.C are treated in Chapter IV.

converge (for $j = 1, 2, \cdots$) and

$$\Gamma_{n_j} \leqq A \qquad \textit{(finite } A\textit{, independent of } j\textit{).} \tag{6.6b}$$

Then the equation (6.2) has a solution $\varphi(y) \subset L_2$.

In order to establish this result we note that by the classical theory the conditions stated in connection with (6.6) and (6.6a) enable us to assert that the equation

$$\int_0^1 K_{n_j}(x, y)\varphi(y)\, dy = f(x) \tag{6.7}$$

has a solution $\varphi_{n_j}(y) \subset L_2$. Clearly

$$\int_0^1 L_x(\xi \mid K_n(x, y))\varphi_n(y)\, dy = L_x(\xi \mid f(x)) \qquad (n = n_j) \tag{6.8}$$

and, inasmuch as the ν^{th} Fourier-coefficient of $\varphi_{n_j}(y)$ is $\lambda_{n_j,\nu} f_{n_j,\nu}$, by Parseval's relation we obtain

$$\int_0^1 \varphi_{n_j}^2(y)\, dy = \Gamma_{n_j}. \tag{6.8a}$$

Thus, in view of (6.6b)

$$\int_0^1 \varphi_{n_j}^2(y)\, dy \leqq A \qquad (j = 1, 2, \cdots). \tag{6.9}$$

In consequence of a known theorem (cf. (3.1), $\cdots$, (3.1b)) it is observed that (6.9) implies that for a suitable choice of the subsequence (n_j) there exists a function $\varphi(y) \subset L_2$ such that

$$\varphi_{n_j}(y) \to \varphi(y) \qquad \text{(weakly)},^7 \tag{6.10}$$

while

$$\int_0^1 \varphi^2(y)\, dy \leqq A. \tag{6.10a}$$

On taking account of (1.9a), (1.9b), of (6.9) and of (6.10), it is concluded that the relation

$$\lim_j \int_0^1 L_x(\xi \mid K_{n_j}(x, y))\varphi_{n_j}(y)\, dy = \int_0^1 L_x(\xi \mid K(x, y))\varphi(y)\, dy = L_x(\xi \mid f(x))$$

is valid by virtue of the known result, stated in connection with (3.5), (3.5a). This establishes Theorem 6.1.

With (6.4) in view and recalling the definition of $\theta_n(x, y \mid \lambda)$, it is deduced that a series (6.6a), if convergent, is representable by the convergent Stieltjes integral

$$\Gamma_{n_j} = \int_{-\infty}^{\infty} \lambda^2\, d_\lambda \int_0^1 f(x) \left(\frac{\partial}{\partial x} \int_0^1 f(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda)\, dy \right) dx. \tag{6.11}$$

On writing

$$(6.11a) \qquad \Gamma_{n_j}(l) = \int_{-l}^{l} \lambda^2 \, d_\lambda \int_0^1 f(x) \left(\frac{\partial}{\partial x} \int_0^1 f(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda) \, dy \right) dx$$

it is deduced that

$$(6.11b) \qquad \Gamma_{n_j}(l) \to \Gamma_{n_j} \qquad (\text{as } l \to \infty)$$

monotonically, inasmuch as the terms of the series (6.6a) are all non-negative. If (6.6b) holds we have

$$(6.11c) \qquad \Gamma_{n_j}(l) \leqq A.$$

Choose the subsequence (n_j) so that the limit (6.4a) exists. It will be convenient to express $\Gamma_{n_j}(l)$ as follows:

$$(6.12) \qquad \begin{aligned} \Gamma_{n_j}(l) &= \int_{-l}^{l} \lambda^2 \, dp_j(\lambda); \qquad p_j(\lambda) = \int_0^1 f(x) q_j(x, \lambda) \, dx; \\ q_j(x, \lambda) &= \frac{\partial}{\partial x} \int_0^1 f(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda) \, dy. \end{aligned}$$

Now

$$(6.12a) \qquad \int_0^1 | q_j(x, \lambda) |^2 \, dx \leqq \int_0^1 f^2(x) \, dx$$

and, hence, the n_j can be so selected that

$$(6.12b) \qquad q_j(x, \lambda) \to q(x, \lambda) = \frac{\partial}{\partial x} \int_0^1 f(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda) \, dy$$

in the weak sense with respect to x;[35]

$$(6.12c) \qquad \int_0^1 q^2(x, \lambda) \, dx \leqq \int_0^1 f^2(x) \, dx.$$

In view of (6.12a) and (6.12b) application of [(3.5), (3.5a)] is possible so as to obtain

$$(6.13) \qquad \lim p_j(\lambda) = p(\lambda) = \int_0^1 f(x) q(x, \lambda) \, dx \qquad (\text{cf. } (6.12b)).$$

We also have

$$(6.14) \qquad | p_j(\lambda) |, \qquad | p(\lambda) |, \qquad \text{Var. } p_j(\lambda), \qquad \text{Var. } p(\lambda) \leqq \int_0^1 f^2(x) \, dx.^{36}$$

In consequence of (6.13) and (6.14) an application of a theorem of *Helly* will result in the relation, valid when (6.6b) holds,

---

[35] Cf., for instance, Trjitzinsky, loc. cit.; [(3.9), (3.9a)]; the latter relations will hold for the more general kernels of the present paper.

[36] This is a consequence of certain formulas in (C), extended by Trjitzinsky, loc. cit. [(3.11)].

$$\lim_j \Gamma_{n_j}(l) = \lim_j \int_{-l}^{l} \lambda^2 \lambda p_j(\lambda) = \int_{-l}^{l} \lambda^2 \, dp(\lambda) \quad [(6.13), (6.12b)], \tag{6.15}$$

provided the n_j are suitably chosen.

By (6.11c) and (6.15)

$$\Gamma(l) = \int_{-l} \lambda^2 \, dp(\lambda) \leqq A, \tag{6.16}$$

where A is independent of l. Now, in view of the manner in which $\Gamma(l)$ has been obtained it follows that $\Gamma(l)$ is a monotone non-decreasing function of l, as $l \to \infty$. Hence, by virtue of (6.16) it is deduced that

$$\lim_l \Gamma(l) = \Gamma = \int_{-\infty}^{\infty} \lambda^2 \, dp(\lambda) = B \leqq A.$$

Thus, *if* (6.6b) *holds then necessarily the integral*

$$\Gamma = \int_{-\infty}^{\infty} \lambda^2 \, d_\lambda \int_0^1 f(x) \left(\frac{\partial}{\partial x} \int_0^1 f(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda) \, dy \right) dx \tag{6.17}$$

exists and $\Gamma \leqq A$, *provided that the subsequence* (n_j) *involved in the definition of* $\Omega(x, y \mid \lambda)$(6.4a) *is suitably chosen.*

The question as to whether there is a converse to this result is of considerable interest, but is at this time left open.

Suppose now merely that the statement with respect to (6.6) *holds.* With $l > 0$ we have

$$\begin{aligned} &\int_0^1 \left[h(x) - \frac{d}{dx} \int_{-l}^{l} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda) \, dy \right]^2 dx \\ &= \left(\int_{-\infty}^{-l} + \int_l^{\infty} \right) d_\lambda \int_0^1 h(x) \left(\frac{\partial}{\partial x} \int_0^1 h(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda) \, dy \right) dx = R(j, l) \end{aligned} \tag{6.18}$$

for all $h(x) \subset L_2$. In fact, (6.18) follows from

$$\begin{aligned} &\int_0^1 \left[h(x) - \sum_\nu{}' \varphi_{n_j,\nu}(x) \int_0^1 h(y) \varphi_{n_j,\nu}(y) \, dy \right]^2 dx \\ &\qquad = \sum_\nu{}'' \left[\int_0^1 h(x) \varphi_{n_j,\nu}(x) \, dx \right]^2, \end{aligned} \tag{6.18a}$$

where the first summation is over the values ν for which $|\lambda_{n_j,\nu}| \leqq l$, while the second summation is over the other values ν. By Parceval's relation

$$\begin{aligned} R(j, 0) &= \int_{-\infty}^{\infty} d_\lambda \int_0^1 h(x) \left(\frac{\partial}{\partial x} \int_0^1 h(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda) \, dy \right) dx \\ &= \int_0^1 h^2(x) \, dx. \end{aligned} \tag{6.19}$$

Now

$$R(j, 0) = A(j, l) + R(j, l), \tag{6.19a}$$

where

$$(6.19b)\qquad A(j, l) = \int_{-l}^{l} d_\lambda \int_0^1 h(x) \left(\frac{\partial}{\partial x} \int_0^1 h(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda)\, dy \right) dx.$$

We shall have

$$(6.20)\qquad \lim_j A(j, l) = A(l) = \int_{-l}^{l} d_\lambda \int_0^1 h(x) \left(\frac{\partial}{\partial x} \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda) \right) dx$$

for a suitable subsequence (n_j); this is established with the aid of a reasoning of the type used in proving (6.15).

By (6.18), if one writes

$$(6.21)\qquad \sigma_j(x \mid l) = h(x) - \frac{d}{dx} \int_{-l}^{l} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda)\, dy,$$

in consequence of (6.19) it is inferred that

$$(6.22)\qquad R(j, l) = \int_0^1 \sigma_j^2(x \mid l)\, dx \leqq R(j, 0) = \int_0^1 h^2(x)\, dx.$$

In view of (6.22) and of [(3.1), $\cdots$ (3.1b)] it is noted that the subsequence (n_j) can be so chosen that there exists a function $\sigma(x \mid l)$ for which

$$(6.22a)\qquad \lim_j \sigma_j(x \mid l) = \sigma(x \mid l) \qquad \text{(weakly)};$$

by (6.22) and (6.22a) one also will have

$$(6.22b)\qquad \lim_j \int_0^1 \sigma_j^2(x \mid l)\, dx = \int_0^1 \sigma^2(x \mid l)\, dx \leqq \int_0^1 h^2(x)\, dx,$$

which is a consequence of [(3.5), (3.5a)].

By virtue of (6.22) and (6.22b)

$$(6.23)\qquad \lim_j R(j, l) = \int_0^1 \sigma^2(x \mid l)\, dx.$$

Moreover, by (6.22a) and (6.21), for $j \to \infty$

$$\int_0^x \sigma_j(x \mid l)\, dx = \int_0^x h(x)\, dx - \int_{-l}^{l} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda)\, dy$$
$$\to \int_0^x \sigma(x \mid l)\, dx = F(x \mid l),$$

where $F(x \mid l)$ is absolutely continuous in x. By a reasoning of the type employed before it is concluded that

$$\lim_j \int_{-l}^{l} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega_{n_j}(x, y \mid \lambda)\, dy = \int_{-l}^{l} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda)\, dy.$$

Thus, for almost all x on (0, 1)

$$(6.23a)\qquad \sigma(x \mid l) = \frac{\partial}{\partial x} F(x \mid l) = h(x) - \frac{d}{dx} \int_{-l}^{l} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda)\, dy$$

and, in view of (6.23),

$$\lim_j R(j, l) = \int_0^1 \left[h(x) - \frac{d}{dx} \int_{-l}^{l} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda)\, dy \right]^2 dx. \tag{6.24}$$

Consequently from (6.19a), (6.20), (6.24) and (6.19) we infer that

$$\begin{aligned} \lim_j R(j, 0) &= \int_{-l}^{l} d_\lambda \int_0^1 h(x) \left(\frac{\partial}{\partial x} \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda)\, dy \right) dx \\ &+ \int_0^1 \left[h(x) - \frac{d}{dx} \int_{-l}^{l} d_\lambda \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda) dy \right]^2 dx = \int_0^1 h^2(x)\, dx. \end{aligned} \tag{6.25}$$

With the aid of certain results, which can be extended to our kernels,[37] it is observed that for the function (6.23a) we have

$$\sigma(x \mid l_\nu) \to \sigma(x) \qquad (\text{weakly in } x;\ l_1 < l_2 < \cdots;\ l_\nu \to \infty), \tag{6.26}$$

where $\sigma(x)$ satisfies

$$\int_0^1 L_x(\xi \mid K(x, y))\sigma(y)\, dy = 0. \tag{6.26a}$$

Clearly

$$\int_0^1 \sigma^2(x \mid l_\nu)\, dx \leqq \int_0^1 h^2(x)\, dx. \tag{6.26b}$$

Turning to (6.25) it is noted that, inasmuch as the first integral displayed is monotone non-decreasing, as $l \to \infty$, both integrals of the second members of (6.25) possess unique limits. Accordingly, put l in (6.25) equal to l_ν and let $\nu \to \infty$. In view of (6.26), of (6.26b) and of [(3.5, (3.5a)]

$$\lim_\nu \int_0^1 \sigma^2(x \mid l_\nu)\, dx = \int_0^1 \sigma^2(x)\, dx.$$

Thus, from (6.25) it is deduced that

$$\begin{aligned} &\int_{-\infty}^{\infty} d_\lambda \int_0^1 h(x) \left(\frac{\partial}{\partial x} \int_0^1 h(y) \frac{\partial}{\partial y} \Omega(x, y \mid \lambda)\, dy \right) dx \\ &\qquad + \int_0^1 \sigma^2(x)\, dx = \int_0^1 h^2(x)\, dx. \end{aligned} \tag{6.27}$$

Since $\sigma(x)$ satisfies (6.26a), from (6.27) it is inferred that the following result will hold.

If for a subsequence (n_j) of (n) each sequence of functions (6.6) forms a comple e set and if the only solution $\subset L_2$ of the equation

$$\int_0^1 L_x(\xi \mid K(x, y))\varphi(y)\, dy = 0 \tag{6.28}$$

---

[37] Cf., for instance, Trjitzinsky, loc. cit. [(3.23), ⋯ (3.23c)].

is $\varphi(y) = 0$ *(almost everywhere) then there exists a closed function* $\Omega(x, y \mid \lambda)$ *(cf.* (6.4a) *and the statement in connection with* (6.5)).

It will be convenient to introduce the following definition.

DEFINITION 6.1. *An operator* $L_x(\xi \mid h(x))$ *will be said to be closed if the relation*

$$L_x(\xi \mid h(x)) = 0 \qquad (h(x) \subset L_2) \tag{6.29}$$

implies $h(x) = 0$ *(almost everywhere on* $(0, 1)$). *The equation*

$$\int_0^1 L_x(\xi \mid K(x, y))\psi(y)\, dy = 0 \tag{6.30}$$

will be said to be closed if the only solution $\subset L_2$ *of* (6.30) *is* $\psi(y) = 0$ *(almost everywhere on* $(0, 1)$.

Suppose now that the equation (6.30) is closed, according to Definition (6.1). With this hypothesis in view suppose, if possible, that there exists an infinite subsequence (m_j) of (n) such that for $m_1, m_2, \cdots$ the following holds.

The equation

$$\int_0^1 L_x(\xi \mid K_{m_j}(x, y))\psi(y)\, dy = 0 \tag{6.31}$$

has a solution $\psi_j(y) \subset L_2$, which is distinct from zero; that is,

$$\int \psi_j^2(y)\, dy \neq 0.$$

Multiplying by suitable constants we arrange to have

$$\int_0^1 L_x(\xi \mid K_{m_j}(x, y))\psi_j(y)\, dy = 0, \qquad \int_0^1 \psi_j^2(y)\, dy = 1 \tag{6.31a}$$

for $j = 1, 2, \cdots$. On taking account of [(3.1), $\cdots$ (3.1b)] the m_j could be so chosen that there exists a function $\psi(y)$ such that

$$\lim_j \psi_j(y) = \psi(y) \qquad \text{(in the weak sense)}. \tag{6.31b}$$

Furthermore, by [(3.5), (3.5a)] we shall have

$$\lim_j \int_0^1 \psi_j^2(y)\, dy = \int_0^1 \psi^2(y)\, dy = 1, \tag{6.31c}$$

corresponding to the aforesaid sequence (m_j).

Now, in view of (1.9a) and (1.9b), together with the second relation (6.31a) and (6.31b), application of [(3.5), (3.5a)] is possible, yielding the result

$$\lim_j \int_0^1 L_x(\xi \mid K_{m_j}(x, y))\psi_j(y)\, dy = \int_0^1 L_x(\xi \mid K(x, y))\psi(y)\, dy = 0, \tag{6.32}$$

where $\psi(y)$ satisfies (6.31c). The equation (6.30) is accordingly not closed, contrary to hypothesis. Hence there exists no infinite subsequence (m_j) of (n) for which (6.31) has a solution $\psi_j(y) \subset L_2$, distinct from zero ($j = 1, 2, \cdots$). Thus, the following has been established.

If the equation (6.30) *is closed* (*Definition* 6.1) *then the equations*

$$\int_0^1 L_x(\xi \mid K_n(x, y))\psi(y)\, dy = 0 \qquad (n = 1, 2, \cdots) \tag{6.33}$$

will be all closed, except at most for a finite number of values n; say, $n(1), n(2), \cdots n(9)$.

Suppose (6.30) is closed; consider an equation (6.33) (for $n \neq n(1), \cdots n(q)$); the equation

$$\int_0^1 K_n(x, y)\psi(y)\, dy = 0 \tag{6.34}$$

will have no solution $\subset L_2$, distinct from zero. In fact, in the contrary case let $\psi_0(y) \subset L_2$ be a solution of (6.34) distinct from zero. Then, by (1.9d)

$$\int_0^1 L_x(\xi \mid K_n(x, y))\psi_0(y)\, dy = 0.$$

This, however, contradicts the fact that the equation (6.33) is closed (for $n \neq n(1), \cdots n(q)$). Thus, with (6.30) closed, every relation

$$\int_0^1 K_n(x, y)\psi(y)\, dy = 0 \qquad (\psi(y) \subset L_2;\, n \neq n(1), \cdots n(q)) \tag{6.34a}$$

will imply that $\psi(y) = 0$ (almost everywhere).

Hence one may replace the result in connection with (6.28) by the following more complete result.

If the equation (6.30) *is closed* (*Definition* 6.1) *then the infinite sequences* (6.6) (*with* $n_i \neq n(1), \cdots n(q)$; cf. (6.33)) *are all complete and there exists at least one closed function* $\Omega(x, y \mid \lambda)$.

We accordingly have the following theorem.

THEOREM 6.2. *The conclusions of Theorem* 6.1 *will hold if the condition stated in connection with* (6.6) *is replaced by the requirement that the equation*

$$\int_0^1 L_x(\xi \mid K(x, y))\psi(y)\, dy = 0$$

be closed in the sense of Definition 6.1.

When the operator L is closed (Definition 6.1) we obtain additional information regarding solubility of the equation

$$\int_0^1 K(x, y)\varphi(y)\, dy = f(x) \tag{6.35}$$

itself. In fact, when L is closed and when the relation

$$\int_0^1 L_x(\xi \mid K(x, y))\varphi(y)\, dy = L_x\left(\xi \,\middle|\, \int_0^1 K(x, y)\varphi(y)\, dy\right)$$

is valid for all $\varphi(y) \subset L_2$, every solution $\varphi(y) \subset L_2$ of the functional equation (6.2) will be a solution of the integral equation (6.35).

UNIVERSITY OF ILLINOIS.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

ON BIANCHI'S PERMUTABILITY THEOREM AND THE THEORY OF W-CONGRUENCES

BY GUIDO FUBINI

(Received May 23, 1939)

1. In the theory of the transformations of many kinds of surfaces and in the theory of rectilinear W-congruences a fundamental rôle is played by Bianchi's permutability theorem. The conviction that it is possible to state a converse theorem has led me to the following calculations and results. In this paper I state also a problem which is not completely solved. We shall say that two surfaces are *congruence-transforms* of each other if the points of the two surfaces are in a one-to-one reciprocal correspondence and the two surfaces are the focal sheets of the congruence generated by the lines joining two corresponding points.

Let the five points x, y, z, t, η be functions of two parameters u, v (that is, their projective homogeneous coördinates x_i, y_i, z_i, t_i, η_i, $[i = 1, 2, 3, 4]$ are functions of u, v). As u, v vary, these points ordinarily generate five surfaces Σ_x, Σ_y, Σ_z, Σ_t, Σ_η.[1] We suppose that *each of* the surfaces Σ_x, Σ_y *is a congruence transform of every one of the surfaces* Σ_z, Σ_t, Σ_η. It may happen that the curved lines enveloped on Σ_x by the tangents (xz), (xt), $(x\eta)$[2] correspond to the lines enveloped on Σ_y by the lines (yz), (yt), $(y\eta)$. In this case every line tangent to the surface Σ_x intersects the corresponding line tangent to Σ_y : the lines (xy) joining corresponding points of Σ_x, Σ_y will pass through a *fixed* point w (whose coördinates w_i are constant). By changing the factors of proportionality we can suppose

$$(1) \qquad y = x + w \quad \text{or} \quad y_i = x_i + w_i \qquad (w_i = \text{const.}; i = 1, 2, 3, 4).$$

We can now change the curvilinear coördinates, and suppose that the tangents (x, z) and (y, z) envelop the lines $v = \text{const.}$ on Σ_x, Σ_y and that the lines (x, t) and (yt) envelop the lines $u = \text{const.}$ Then the tangents (x, z) and (y, z) are identical with the tangents (x, x_u) and (y, y_u); and their intersection z is identical with the point

$$(2) \qquad z = x_u = y_u \qquad \left[z_i = \frac{\partial x_i}{\partial u} = \frac{\partial y_i}{\partial u}; (i = 1, 2, 3, 4)\right].$$

[1] Corresponding points on these surfaces are the points defined by the same values of u, v. Σ_x may not be a congruence transform of Σ_y.

[2] By (xz) or, by (x, z), for instance, we indicate the line joining the points x and z; this line touches the surfaces Σ_x and Σ_z at these points.

In the same way we prove that we can suppose

$$t = x_v = y_v. \tag{3}$$

The line through z and t also passes through η; therefore we can find a function $\rho(u, v)$ such that

$$\eta = z + \rho t = x_u + \rho x_v = y_u + \rho y_v. \tag{4}$$

From our hypothesis it follows that the plane tangent at z to the surface Σ_z, the plane tangent at t to Σ_t, the plane tangent at η to the surface Σ_η all pass through x and y, that is, through x and w. Therefore z_u, z_v must be linear combinations of z, x, w; t_u, t_v linear combinations of t, x, w; and η_u, η_v linear combinations of η, x, w. We can find eighteen functions $a, b, c, A, B, C, p, q, \cdots$ such that

$$\begin{aligned} z_u &= az + bx + cw, & t_u &= pt + qx + rw, & \eta_u &= l\eta + mx + nw, \\ z_v &= Az + Bx + Cw, & t_v &= Pt + Qx + Rw, & \eta_v &= L\eta + Mx + Nw. \end{aligned} \tag{5}$$

By use of (4) we deduce:

$$\rho_u = a - p, \qquad \rho_v = A - P; \tag{6}$$

and therefore $(a - p)_v = (A - P)_u$. From the identities $z_v = x_{uv} = t_u$ we deduce:

$$A = p = 0, \qquad q = B, \qquad r = C.$$

From the identities $(z_u)_v = (z_v)_u$ and $(t_u)_v = (t_v)_u$ we deduce:

$$b = 0, \quad B = a_v, \quad Q = 0, \quad P_u = q, \quad C_u = c_v + aC, \quad r_v = R_u + rP,$$

and from (6) we obtain

$$b = Q = A = p = q = B = 0, \qquad a_v = 0, \qquad P_u = 0, \qquad r = C,$$

$$r_u = c_v + ar, \qquad r_v = R_u + rP.$$

The equations (5) become:

$$x_{uu} = z_u = az + cw = ax_u + cw$$

$$x_{uv} = z_v = t_u = rw$$

$$x_{vv} = t_v = \quad = Px_v + Rw,$$

where a is a function only of u, and P is a function of v. We can change the parameters of the coördinate lines $u =$ const. and $v =$ const. in such a way that $a = P = 0$; and the foregoing equations become

$$x_{uu} = cw, \qquad x_{uv} = rw, \qquad x_{vv} = Rw.$$

From the conditions of integrability it follows that we can find a function φ such that

$$c = \varphi_{uu}, \qquad r = \varphi_{uv}, \qquad R = \varphi_{vv}.$$

Therefore the foregoing equations become:

$$(x - \varphi w)_{uu} = (x - \varphi w)_{uv} = (x - \varphi w)_{vv} = 0.$$

Thence

$$x_i = \varphi w_i + a_i u + b_i v + c_i \qquad (w_i, a_i, b_i, c_i = \text{const.}) \tag{7}$$

or

$$x = \varphi w + ua + vb + c, \tag{7'}$$

where w, a, b, c are fixed points. From (6) we get $\rho = \text{const.}$, and, if we put $\eta = z + \rho t$, where ρ is an *arbitrary* constant, we find that *all the* ∞^1 *surfaces* Σ_η (and only these surfaces) *are congruence-transforms of both of the surfaces* Σ_x, Σ_y. We indicate, as usual, by (x, η), (x, η_u) $\cdots$ also the coördinates of the line joining the points x and η, of the line joining x and η_u, $\cdots$. From $(x, \eta)_u = (x_u, \eta) + (x, \eta_u)$, $\cdots$, we deduce:

$$(x, \eta)_u = \rho(x_u, x_v) + (\varphi_{uu} + \rho\varphi_{uv})(x, w),$$

$$(x, \eta)_v = (x_v, x_u) + (\varphi_{uv} + \rho\varphi_{vv})(x, w), \cdots.$$

It is easy to see that (x, η), $(x, \eta)_u$, $(x, \eta)_v$, $(x, \eta)_{uu}$, $(x, \eta)_{uv}$, $(x, \eta)_{vv}$ are *not* linearly independent, because they are linear combinations of (x, η), (x_u, x_v), (x, w), (x_u, w), (x_v, w). Hence: *the congruences generated by the lines* (x, η) *are all W-congruences*; in the same way we demonstrate that also *the lines* (y, η) *generate W congruences. In this case* we have therefore completely demonstrated *the converse of Bianchi's permutability theorem.* And in this case we remark: *Not only* Σ_x *and* Σ_y *are congruence-transforms of all the* ∞^1 *surfaces* Σ_η, *but also the* ∞^1 *surfaces generated by the points* $x + \sigma w$ *(where* σ *is an arbitrary constant) are congruence-transforms of all the surfaces* Σ_η.

2. Let us now suppose that *only* two of the tangents (x, z), (x, t), (x, η) are the corresponding tangents of (y, z), (y, t), (y, η): for instance we suppose that (x, z) and (y, z) [(x, t) and (y, t)] are corresponding tangents, but that (x, η) does *not* correspond to (y, η). Let us change the parameters u, v so that the lines $v = \text{const.}$ are the lines enveloped on Σ_x and Σ_y by the tangents (x, x) and (y, z) and that the lines $u = \text{const.}$ are the lines enveloped by (x, t) and (y, t). We can find twenty functions a, b, $\cdots$ of u, v such that

$$x_u = ax + bz \qquad y_u = Ay + Bz \tag{8}$$

$$x_v = px + qt \qquad y_v = Py + Qt \tag{9}$$

$$\begin{aligned} z_u &= hz + kx + ly & t_u &= ct + ex + fy \\ z_v &= Hz + Kx + Ly & t_v &= Ct + Ex + Fy. \end{aligned} \tag{10}$$

Obviously[3]

$$b \neq 0, \qquad q \neq 0, \qquad B \neq 0, \qquad Q \neq 0. \tag{11}$$

[3] If, for instance, $b = 0$, then $x_u = ax$; and the point x could not generate a surface.

Since the point η belongs to the line passing through z and t, we can write

$$\eta = z + \rho t, \tag{12}$$

where ρ is a function of u, v. Also η_u and η_v must be linear combinations of η, x, y. If we calculate η_u and η_v, we deduce by means of (10):

$$\frac{\rho_u}{\rho} = h - c, \qquad \frac{\rho_v}{\rho} = H - C,$$

$$(h - c)_v = (H - C)_u.$$

From the last equations it follows that we can multiply the homogeneous coördinates of z by the same factor, and the coördinates of t by another factor, so that

$$h = c, \qquad H = C,$$

(we could also make h and c equal to zero). Therefore $\rho = \text{const.}$

If ρ is an arbitrary constant, η_u and η_v are always linear combinations of η, x, y; and all the ∞^1 surfaces (depending *only* on one constant ρ) are congruence-transforms of Σ_x and Σ_y. The lines $(x, \eta) = (x, z + \rho t)$ envelop on Σ_x the curved lines defined by the equation

$$\frac{du}{b} = \frac{dv}{q:\rho}$$

and the lines (y, η) envelop on Σ_y the curves

$$\frac{du}{B} = \frac{dv}{Q:\rho}.$$

Because of our hypothesis, these curves do not correspond; and therefore $b:q \neq B:Q$. If we give to ρ two values ρ and ρ_1 ($\rho \neq \rho_1$; $\rho, \rho_1 \neq 0, \infty$) we obtain two points η and η_1, which generate two surfaces Σ_η and Σ_{η_1}, both of which are congruence-transforms of Σ_x and Σ_y. And the tangent (x, η) does not correspond to (y, η); the tangent (x, η_1) does not correspond to (y, η_1). Therefore by changing η and η_1 into z, t, we can suppose that (x, z) does not correspond to (y, z) and (x, t) does not correspond to (y, t).

3. Let us now suppose that the tangents (x, z) to S_x and (yz) to S_y are not corresponding tangents. We can change parameters so that the lines (x, z) envelop on Σ_x the lines $v = \text{const.}$, and the lines (y, z) envelop on Σ_y the lines $u = \text{const.}$ Therefore there exist four functions a, b, A, B of u, v, such that

$$x_u = ax + bz, \qquad y_v = Ay + Bz. \tag{13}$$

Since the plane tangent at z to Σ_z is the plane of the points x, y, z, we can find six functions k, s, w, $\cdots$ so that

$$z_u = kz + sx + wy, \qquad z_v = Kz + Wx + Sy. \tag{14}$$

The tangent plane at x to Σ_x [at y to Σ_y] is the plane of the points x, z, t [of the points y, z, t]. Therefore x_v is a linear combination of x, z, t, and y_u of y, z, t. We can write:

$$(13') \qquad x_v = cx + fz + gt, \qquad y_u = Cy + Fz + Gt,$$

where $c, f \cdots$ are new functions of u, v.

We can also find for t two equations which are analogous to the equations (14) for z:

$$(14') \qquad t_u = ht + mx + ny, \qquad t_v = Ht + Nx + My.$$

Instead of (13′) we can write the equations

$$(13'') \quad bx_v - fx_u = (bc - af)x + bgt; \quad By_u - Fy_v = (BC - AF)y + BGT.$$

If also the tangents (xt) and (yt) are not corresponding tangents, the difference $bB - fF$ is not equal to zero; and we can choose the factors $1:\rho$ and $1:\sigma$ so that

$$dU = \frac{1}{\rho}\,\frac{Fdu + Bdv}{bB - fF}, \qquad dV = \frac{1}{\sigma}\,\frac{fdv + bdu}{bB - fF}$$

(whence $du = -f\rho\, dU + B\sigma\, dV$; $dv = b\rho\, dU - F\sigma\, dV$) become exact differentials. We can substitute for the equations (13′) the equivalent equations (13″), which we have now transformed into

$$(13''') \qquad \frac{\partial x}{\partial U} = \rho(bc - af)\,x + \rho bgt; \frac{\partial y}{\partial V} = \sigma(BC - AF)y + \sigma BGt,$$

and these equations are completely analogous to the (13). Therefore there is no substantial difference between the points z and t, or between the equations (13), (13′). Also in this case we can write:

$$\eta = z + \rho t;$$

and the η_u, η_v must be linear combinations of η, x, y. We deduce

$$\frac{\rho_u}{\rho} = k - h; \qquad \frac{\rho_v}{\rho} = K - H; \qquad (k - h)_v = (K - H)_u$$

and we can multiply the homogeneous coördinates of t and z by such factors that

$$(15) \qquad K = H; \qquad k = h = 0.$$

Then $\rho =$ const.; and, by giving to ρ an arbitrary constant value, we *obtain* ∞^1 *surfaces* Σ_η *which are all congruence transforms of* Σ_x, Σ_y; and conversely, each surface, congruence transform of Σ_x and Σ_y, is one of the preceding surfaces Σ_η. From (13), (14), (13′), (14′), we can deduce the values of the derivatives of $x_u, x_v, y_u, y_v, z_u, z_v, t_u, t_v$ as linear combinations of x, y, z, t. The conditions of integrability

$$\frac{\partial}{\partial v}\, x_u = \frac{\partial}{\partial u}\, x_v, \qquad \frac{\partial}{\partial v}\, y_u = \frac{\partial}{\partial u}\, y_v, \qquad \frac{\partial}{\partial u}\, z_v = \frac{\partial}{\partial v}\, z_u, \qquad \frac{\partial}{\partial u}\, t_v = \frac{\partial}{\partial v}\, t_u$$

must be transformed as identities, because the points x, y, z, t are linearly independent. Let us now compare the coefficients of y in the first of these equations, the coefficients of x in the second, the coefficients of t in the third, and the coefficients of z in the fourth. We get, from (15):

$$bS = fw + gn \tag{16}$$

$$Bs = FW + GN \tag{16'}$$

$$sg = SG \tag{16''}$$

$$mf + nB = MF + Nb. \tag{16'''}$$

By comparing the coefficients of z (of t) in the third (the fourth) of the preceding identities, we have:

$$sf + wB = SF + Wb + H_u, \qquad mg = MG + H_u,$$

whence, by eliminating H_u:

$$j = bW - fs - GM = Bw - FS - gm. \tag{17}$$

(We have called j the value of the second and third members.) From these identities it follows:

If one of the congruences generated by one of the lines (x, z), (x, t), (x, η), (y, z), (y, t), (y, η) *is a W-congruence, these congruences are all W-congruences* (and we have precisely the case of Bianchi's permutability theorem).

In order to demonstrate this proposition, we begin by remarking that, by changing notations, we can indicate by Σ_z, Σ_t two surfaces arbitrarily chosen among the ∞^1 surfaces Σ_η. Therefore it is sufficient to prove:

If the congruence generated by (x, z) *is W, the three congruences generated by one of the lines* (x, t), (y, z), (y, t) *are also W-congruences.* The congruence $K_{xz} = K_{zx}$ generated by the lines (x, z) is W if and only if the coördinates (x, z) of the lines (x, z) and their derivatives of first and of second order are not linearly independent. By means of the preceding equations we find that all these derivatives are linear combinations of the coördinates of the six lines (x, z), (y, z), (x, t), (y, t), (x, y), (t, z). The congruence K_{xz} is W if and only if the determinant Δ_{xz} of the coefficients of these combinations is equal to zero. We find

$$(x, z)_u = \frac{\partial}{\partial u}(x, z) = (x_u, z) + (x, z_u) = a(x, z) + w(x, y),$$

$$(x, z)_v = (c + H)(x, z) + g(t, z) + S(x, y).$$

Disregarding the terms in (x, z), (x, y), (t, z), we find also

$$(x, z)_{uu} = w[b(z, y) + G(x, t)] + \cdots,$$

$$(x, z)_{uv} = w[b(z, y) + g(t, y)] + \cdots,$$

$$(x, z)_{vv} = (Sf - Mg)(zy) + 2Sg(t, y) - gW(x, t) + \cdots.$$

Therefore

$$\pm\Delta_{xz} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ a & w & 0 & 0 & 0 & 0 \\ c+H & S & g & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & wb & wG & 0 \\ \cdots & \cdots & \cdots & wf & 0 & wg \\ \cdots & \cdots & \cdots & Sf - Mg & -gW & 2Sg \end{vmatrix}$$

$$= w^3g \begin{vmatrix} b & G & 0 \\ f & 0 & g \\ Sf - Mg & -gW & 2Sg \end{vmatrix} = w^3g^2 \begin{vmatrix} b & G & 0 \\ f & 0 & g \\ -M & -W & S \end{vmatrix}$$

$$= g^2w^3(bgW - GfS - gGM).$$

By means of (16″) and (17) we deduce:

$$\pm\Delta_{xz} = g^3w^3j. \tag{18}$$

By analogous calculations we find

$$(y, z)_u = C(y, z) + G(t, z) + s(y, x),$$

$$(y, z)_v = (A + H)(y, z) + W(y, x),$$

and, by *disregarding* the terms in (y, z), (y, x), (t, z):

$$(yz)_{uu} = (Gm - sF)(x, z) + 2Gs(t, x) + Gw(t, y) + \cdots,$$

$$(yz)_{uv} = G[N(x, z) + W(t, x) + S(t, y)] - s[g(t, y) + B(x, z)] + \cdots,$$

$$(yz)_{vv} = -W[g(t, y) + B(x, z)] + \cdots,$$

and we find immediately that the determinant Δ_{yz} of the coefficients of these linear combinations is given by

$$\pm\Delta_{yz} = G^2W^2 \begin{vmatrix} Gw & Gm - sF & 2Gs \\ S & N & W \\ g & B & 0 \end{vmatrix}$$

$$= G^2W^2\, g(GmW - 2sGN - sFW) - B(GwW - 2sSG)$$

$$= G^3W^2\,(gmW - 2GSN - SFW - BwW + 2sSB)\ {}^{4}$$

$$= G^3W^3\,(gm + SF - Bw),\ {}^{5}$$

and therefore

$$\pm\Delta_{yz} = G^3W^3j. \tag{18'}$$

[4] By means of (16″).
[5] By means of (16′).

Let us turn now to (xt). We find

$$(x, t)_u = a(x, t) + b(z, t) + n(x, y),$$

$$(x, t)_v = (c + H)(xt) + f(z, t) + M(x, y),$$

and, by *disregarding* the terms in (x, t), (z, t), (x, y):

$$(x, t)_{uu} = b\{-m(x, z) - n(y, z) - w(t, y)\} + n\{b(z, y) + F(x, z)\} + \cdots,$$

$$(x, t)_{uv} = b\{-N(x, z) - M(y, z) - S(t, y)\} + n\{B(x, z) + f(z, y) + g(t, y)\} + \cdots,$$

$$(x, t)_{vv} = f\{-N(x, z) - M(y, z) - S(t, y)\} + M\{B(x, z) + f(z, y) + g(t, y)\} + \cdots.$$

Therefore the determinant of the coefficients of these linear combinations is given by

$$\pm\Delta_{xt} = \begin{vmatrix} 1 & 0 & 0 \\ a & b & n \\ c + H & f & M \end{vmatrix} \cdot \begin{vmatrix} b & n \\ f & M \end{vmatrix} \cdot \left\{ b \begin{vmatrix} m & n & w \\ N & M & S \\ B & -f & g \end{vmatrix} - n \begin{vmatrix} F & -b & 0 \\ N & M & S \\ B & -f & g \end{vmatrix} \right\}.$$

In the right-hand member $-n$ is multiplied by

$$\begin{vmatrix} F & -b & 0 \\ N & M & S \\ B & -f & g \end{vmatrix} = g(FM + bN) + FfS - B(bS)$$

$$= g(mf + nB) + FfS - B(fw + gn)$$

(by means of (16‴) and (16). Therefore $-n$ is multiplied by $-fj$. The coefficient of b is

$$\begin{vmatrix} m & n & w \\ N & M & S \\ B & -f & g \end{vmatrix} = M(mg - wB + SF) + S(nB - MF + mf) - nNg - wNf$$

$$= -Mj + N(Sb - ng - wf) = -Mj$$

(by means of (16‴) and (16)). Therefore

(18″) $$\pm\Delta_{xt} = (bM - nf)^3 j.$$

In the same way we find:

$$(y, t)_u = C(y, t) + F(z, t) + m(y, x),$$

$$(y, t)_v = (A + H)(y, t) + B(z, t) + N(y, x),$$

and, by *disregarding* the terms in (y, t), (z, t), (y, x):

$$(y, t)_{uu} = -F[m(x, z) + n(y, z) + s(t, x)] - m[b(z, y) + F(x, z) + G(x, t)] + \cdots,$$

$$(y, t)_{uv} = -F[N(x, z) + M(y, z) + W(t, x)] - m[f(z, y) + B(x, z)] + \cdots,$$

$$(y, t)_{vv} = -B[N(x, z) + M(y, z) + W(t, x)] - N[f(z, y) + B(x, z)] + \cdots.$$

Therefore the determinants Δ_{yt} of the coefficients is given by

$$\pm\Delta_{yt} = \begin{vmatrix} 1 & 0 & 0 \\ C & F & m \\ A+H & B & N \end{vmatrix} (FN - Bm) \left\{ -m \begin{vmatrix} b & F & G \\ -M & N & -W \\ f & B & 0 \end{vmatrix} - F \begin{vmatrix} -n & m & -s \\ -M & N & -W \\ f & B & 0 \end{vmatrix} \right\}$$

$$= (FN - mB)^2[-bWB + f(FW + NG) + GMB]m + [nBW + fmW - sfN - M(sB)]F$$

$$= (FN - mB)^2\{mB(-bW + fs + GM) + F(nBW + fmW - sfN - M[FW + GN])\}\ ^{6}$$

$$= (FN - mB)^2\{(mB - FN)(-bW + fs + GM) + W(-bFN + FnB + Ffm - MF^2\}$$

$$= (FN - mB)^3(bW - fs - GM).\ ^{7}$$

Therefore

$$(18''') \qquad \pm\Delta_{yt} = (FN - mB)^3 j.$$

Let us now remark that

A) $wg \neq 0$. For, if $g = 0$, the tangents (x, x_u) and (x, x_v) would coincide. If $w = 0$, x_{uu} would be a linear combination of x, x_u. And in this case the lines $v =$ const. on Σ_x would be straight lines, and their tangents (x, x_u) would not generate a congruence.

B) $WG \neq 0$, because the tangents (y, y_u) and (y, y_v) cannot coincide, and because y_{vv} cannot be a linear combination of y, y_v.

C) $bM - nf \neq 0$. Together with (13‴) we can write

$$\frac{\partial t}{\partial U} = \rho\left(-f\frac{\partial t}{\partial u} + b\frac{\partial t}{\partial v}\right) = \rho\{H(b - f)t + (bN - fm)x + (bM - fn)y\}.$$

If $bM - nf = 0$, the derivatives $\partial^2 x/\partial U^2$ would be a linear combination of x and $\partial x/\partial U$.

D) $FN - mB \neq 0$. Together with (13‴) we can write

$$\frac{\partial t}{\partial V} = \sigma\left(B\frac{\partial t}{\partial u} - F\frac{\partial t}{\partial v}\right) = \sigma\{H(B - F)t + (Bm - FN)x + (Bn - FM)y\}.$$

[6] By means of (16′).
[7] By means of (16‴).

If $FN - mB = 0$, the derivative $\partial^2 y/\partial V^2$ would be a linear combination of y and $\partial y/\partial V$.

From these remarks and from (18), (18′), (18″), (18‴) it follows that, if one of the determinants Δ_{xz}, Δ_{yz}, Δ_{zt}, Δ_{yt} is equal to zero, then $j = 0$, and all these determinants are equal to zero; which demonstrates precisely the stated theorem. But this theorem does not prove that always $j = 0$, i.e. that the congruences are all W-congruences. We must now remember that in Bianchi's theorem there are *two* systems of ∞^1 surfaces, *and that every surface of a system is a congruence-transform of every surface of the other system.* In this case we can complete the preceding result, and *demonstrate precisely that all congruences are W-congruences.* We must suppose we can find a function $\sigma(u, v)$ such that the surface generated by the point $x + \sigma y$ will be congruence-transforms of Σ_z, Σ_t. This is true only if $d(x + \sigma y)$ is a linear combination of $x + \sigma y$, z, t. But

$$d(x + \sigma y) = x(a\,du + c\,dv) + \sigma\left[C\,du + A\,dv + \frac{d\sigma}{\sigma}\right]y + \cdots,$$

where we have not written the terms in z or t. Our condition will be satisfied if and only if

$$a\,du + c\,dv = C\,du + A\,dv + \frac{d\sigma}{\sigma}$$

that is, if $(C - a)\,du + (A - c)\,dv$ is an exact differential (and σ will depend *only* on *one* arbitrary constant). We can therefore find σ, if and only if

$$\frac{\partial(A - c)}{\partial u} = \frac{\partial(C - a)}{\partial v}.$$

In the condition of integrability $\partial x_u/\partial v = \partial x_v/\partial u$, let us write the term in x; and in the identity $\partial y_u/\partial v = \partial y_v/\partial u$, let us write the term in y. We find

$$\frac{\partial a}{\partial v} - \frac{\partial c}{\partial u} = fs + gm - bW, \qquad \frac{\partial C}{\partial v} - \frac{\partial A}{\partial u} = Bw - FS - GM.$$

These two derivatives will be equal to each other if and only if

$$fs + gm - bW = Bw - FS - GM \text{ or } fs + GM - bW = Bw - FS - gm.$$

By comparing with the identity (17) we get $j = 0$; which proves the theorem quoted.

4. We solve now another problem. *When does it happen that two surfaces Σ_x, Σ_y, congruence transforms of an identical surface Σ_z, are also congruence-transforms of at least two other surfaces Σ_t, and consequently of ∞^1 surfaces Σ_t?*

After changing notations, let us suppose that u, v are asymptotic parameters on Σ_z and that z are projective *non*-homogeneous coördinates. It is known[8] that we can determine seven functions θ, β, γ, A, B, A', B' of u, v such that

[8] Fubini e Čech, *Geometria proiettivo differenziale* (Bologna, Zanichelli), vol. I, pp. 89, 90, 243, 244.

$$z_{uu} = \theta_u z_u + \beta x_v \qquad z_{vv} = \theta_v z_v + \gamma z_u$$

$$x = \mu z + 2(A z_u + B z_v) \qquad y = \mu' z + 2(A' z_u + B' z_v)$$

$$\xi = \lambda \zeta + 2(A \zeta_u - B \zeta_v) \qquad \eta = \lambda' \zeta + 2(A' \zeta_u - B' \zeta_v),$$

where ζ, ξ, η are the coordinates of the planes tangent at z, x, y to the surfaces Σ_z, Σ_x, Σ_y, and where

$$\mu = -A_u - A\theta_u + A\,\frac{B_u + A\beta}{B} + \left(-B_v - B\theta_v + B\,\frac{A_v + B\gamma}{A}\right),$$

$$\lambda = -A_u - A\theta_u + A\,\frac{B_u + A\beta}{B} + \left(B_v + B\theta_v - B\,\frac{A_v + B\gamma}{A}\right),$$

and μ', λ' are given by the analogous equations, obtained by changing A, B, λ, μ into A', B', λ', μ'.

The points t lie on the intersection of the planes ξ, η, which is the straight line joining the points z and

$$4\Delta z_{uv} + (\lambda' x - \lambda y) \qquad [\Delta = A'B - AB'].$$

Therefore we can write

$$t = Rz + 4\Delta z_{uv} + (\lambda' x - \lambda y).$$

We must determine R so that the plane tangent at t to $\sum_t$ shall be the plane passing through t, x, y; we require also that dt be a linear combination of t, x, y. We remark that

$$dz_{uv} = z_{uv} d\theta + \omega_1 z_u + \omega_2 z_v,$$

$$d(\lambda' x - \lambda y) = \omega_3 z + \omega_4 z_{uv} + \omega_5 x + \omega_6 y,$$

where

$$\omega_1 = (\theta_{uv} + \beta\gamma)\,du + \pi_{22}\,dv \qquad (\pi_{22} = \gamma_u + \gamma\theta_u)$$

$$\omega_2 = \pi_{11}\,du + (\theta_{uv} + \beta\gamma)\,dv \qquad (\pi_{11} = \beta_v + \beta\theta_v)$$

$$\omega_3 = \left\{\lambda'\left(\mu_u - \frac{B_u + A\beta}{B}\,\mu\right) - \lambda\left(\mu'_u - \frac{B'_u + A'\beta}{B'}\,\mu'\right)\right\} du$$
$$+ \left\{\lambda'\left(\mu_v - \frac{A_v + B\gamma}{A}\,\mu\right) - \lambda\left(\mu'_v - \frac{A'_v + B'\gamma}{A'}\,\mu'\right)\right\} dv,$$

$$\omega_4 = 2(\lambda' B - \lambda B')\,du + (\lambda' A - \lambda A')\,dv$$

and where we need not calculate ω_5 and ω_6. Therefore

$$4\Delta(dt - \omega_5 x - \omega_6 y) - (\omega_4 + 4\Delta + 4\Delta\,d\theta)t = 4\Delta(dR + \omega_3)$$
$$- R(\omega_4 + 4\Delta + 4\Delta\,d\theta)z + 4\Delta(R\,du + 4\Delta\omega_1)z_u + 4\Delta(R\,dv + 4\Delta\omega_2)z_v$$

is a linear combination *only* of z, z_u, z_v (and *not* of z_{uv}). Our condition is satisfied only if this expression and x, y are not linearly independent,[9] i.e., if

$$dR = (\Omega + \Omega')R - \omega_3 - 2\omega_1(B'\mu - B\mu') + 2\omega_2(A'\mu - A\mu') \tag{19}$$

where

$$\Omega = \frac{B_u + A\beta}{B}du + \frac{A_v + B\gamma}{A}dv, \tag{19'}$$

and Ω' is given by the analogous equation deduced by changing A, B, into A', B'. According to our hypothesis that R has *at least two* values, the condition of the integrability of (19) must be identically satisfied; and therefore R has ∞^1 values, depending only on *one* arbitrary constant. The *first* condition of integrability is

$$\Omega + \Omega' = \text{an exact differential.}$$

We find again a theorem already proved in the preceding pages.

If one of the congruences (x, z) *and* (y, z) *is* W (that is, if Ω or Ω' is an exact differential), *the other congruence also is* W. By multiplying A, B (and consequently μ) by the same factor r, the point x does not change; and the point y does not change if we multiply A', B' by the same factor r'. If $\Omega + \Omega'$ is exact, we can obtain (by a suitable choice of these factors) that $\Omega + \Omega' = 0$, and also that

$$\begin{aligned} &A_v + B\gamma = A'_v + B'\gamma = 0; \qquad \frac{B_u + A\beta}{B} = \varphi; \\ &\frac{B'_u + A'\beta}{B'} = \varphi', \quad \text{where} \quad \varphi + \varphi' = 0; \end{aligned} \tag{20}$$

and, by supposing that these equations are satisfied, we find that the second condition of integrability is

$$\begin{aligned} 0 = 2\varphi_v[\mu\mu' &+ \mu(B'_v + B'\theta_v) + \mu'(B_v + B\theta_v) \\ &+ \varphi(A\mu' - A'\mu) - (B\mu'_v + B'\mu_v) + (A\mu'_u + A'\mu_u)]. \end{aligned} \tag{21}$$

[9] Our condition is equivalent to the equation

$$0 = \begin{vmatrix} 4\Delta(dR + \omega_3) - R(\omega_4 + 4\,d\Delta + 4\Delta\,d\theta) & 4\Delta(R\,du + 4\Delta\omega_1) & 4\Delta(R\,dv + 4\Delta\omega_2) \\ \mu & 2A & 2B \\ \mu' & 2A' & 2B' \end{vmatrix}.$$

By expanding we find that in the right-hand member the coefficient of dR is $-(4\Delta)^2$ and the coefficient of R is

$$4\Delta(\omega_4 + 4d\Delta + 4\Delta\,d\theta) - 4\Delta(2\mu B' - 2B\mu')\,du + 4\Delta(2\mu A' - 2\mu' A)\,dv = (4\Delta)^2(\Omega + \Omega').$$

From this remark we can deduce (19) immediately.

This can be seen by observing that by means of (20), equation (19) becomes equivalent to

$$dR = \varphi(\lambda'\mu + \lambda\mu')\,du + \{(\lambda\mu'_u - \lambda'\mu_u)\,du + (\lambda\mu'_v - \lambda'\mu_v)\,dv\} + 2\{(\mu P' - \mu' P)\,du + (\mu Q' - \mu' Q)\,dv\}$$

where

$$P = \pi_{11}A - B(\theta_{uv} + \beta\gamma); \qquad Q = A(\theta_{uv} + \beta\gamma) - \beta\pi_{22},$$

and P', Q' are given by quite analogous equations. The conditions of integrability of the projective geometry of the surfaces (loc. cit. p. 94) are equivalent to

$$(\theta_{uv} + \beta\gamma)_u + \beta\pi_{22} - \frac{\partial\pi_{11}}{\partial v} = (\theta_{uv} + \beta\gamma)\theta_u$$

$$(\theta_{uv} + \beta\gamma)_v + \gamma\pi_{11} - \frac{\partial\pi_{22}}{\partial u} = (\theta_{uv} + \beta\gamma)\theta_v.$$

From these equations and from (20) we deduce:

$$P = -(B_v + B\theta_v)_v + (B\varphi)_v + B\varphi\theta_v\,; \qquad Q = (A_u + A\theta_u)_v + B\varphi\gamma,$$
$$Q_u - P_v = -\mu(\theta_{uv} + \beta\gamma) + \varphi Q,$$

and the analogous equations for P', Q', A', B', $\varphi' = -\varphi$. From $\lambda = \mu + 2(B_v + B\theta_v)$ and the analogous identity for λ', μ' we get:

$$\frac{\partial}{\partial v}(\lambda\mu'_u - \lambda'\mu_u) - \frac{\partial}{\partial u}(\lambda\mu'_v - \lambda'\mu_v) = \lambda_v\mu'_u - \lambda'_v\mu_u - \lambda_u\mu'_v + \lambda'_u\mu_v$$
$$= 2(\mu + B_v + B\theta_v)_v\mu'_u - 2(\mu' + B'_v + B'\theta_v)\mu_u + 2(B'_v + B'\theta_v)\mu_v - 2(B_v + B\theta_v)\mu'_v;$$

$$\frac{\partial}{\partial v}[\varphi(\lambda\mu' + \lambda'\mu)] = 2\frac{\partial}{\partial v}\{\varphi[\mu\mu' + \mu(B'_v + B'\theta_v) + \mu'(B_v + B\theta_v)]\}$$
$$= 2\varphi_v[\mu\mu' + \mu(B'_v + B'\theta_v) + \mu'(B_v + B\theta_v)] + 2\varphi\{\mu'(-A_u - A\theta_u + A\varphi)_v + \mu(-A'_u + A'\theta_u - A'\varphi)_v + (B_v + B\theta_v)\mu'_v + (B'_v + B'\theta_v)\mu_v\};$$

$$2\left\{\frac{\partial}{\partial v}(\mu P' - \mu' P) - \frac{\partial}{\partial u}(\mu Q' - \mu' Q)\right\}$$
$$= 2[\mu_v P' - \mu'_v P - \mu_u Q' + \mu'_u Q + \varphi(\mu Q' + \mu' Q)].$$

By summing we find that dR is an exact differential only if

$$0 = 2\varphi_v\{\mu\mu' + \mu(B'_v + B'\theta_v) + \mu'(B_v + B\theta_v)\} + 2\varphi\mu'[Q + (A\varphi - A\theta_u - A_u)_v]$$
$$+ 2\varphi\mu[Q' - (A'_u + A'\theta_u + A'\varphi]_v + 2\mu_v[P' + \varphi(B'_v + B'\theta_v) + (B'_v + B'\theta_v)_u]$$
$$- 2\mu'_v[P - \varphi(B_v + B\theta_v) + (B_v + B\theta_v)_u] - 2\mu_u[Q' + (\mu' + B'_v + B'\theta_v)_v]$$
$$+ 2\mu'_u[Q + (\mu + B_v + B\theta_v)_v],$$

and this equation is equivalent to (21).

I have not been able to find the *geometrical meaning* of (21) nor to study the system of equations (20), (21). We can only remark that equation (21) is identically satisfied if $\varphi = 0$ [and also if only $\varphi_v = 0$; but in this case we can, by a suitable choice of the factors r and r', make $\varphi = 0$]. But if $\varphi = 0$, both of the congruences generated by the lines (x, z) and (y, z) are W-congruences. *This remark gives therefore a new proof of Bianchi's permutability theorem.*

5. We can also generalize some of the preceding results. Let us suppose that Σ_x is a congruence-transform of ∞^1 surfaces Σ_t, that the points t corresponding to the same point x lie on a straight line, and that the corresponding planes τ tangent at t to Σ_t form a pencil of planes. (These conditions are satisfied if the Σ_t, as in the preceding section, are congruence transforms also of another surface Σ_y.) Changing notations, we suppose that u, v are asymptotic parameters on Σ_x, and we calculate β, γ, θ not for Σ_t, but for Σ_x. We find that we can put

$$t = \mu x + 2(Ax_u + Bx_v),$$
$$\tau = \lambda\xi + 2(A\xi_u - B\xi_v),$$

where

$$\mu = -A_u - A\theta_u - B_v - B\theta_v + A\,\frac{B_u + A\beta}{B} + B\,\frac{A_v + B\gamma}{A},$$

$$\lambda = -A_u - A\theta_u + B_v + B\theta_v + A\,\frac{B_u + A\beta}{B} - B\,\frac{A_v + B\gamma}{A}.$$

In these equations A, B depend on one constant c (because we have ∞^1 surfaces Σ_t) and ξ are the coördinates of the plane tangent at x to Σ_x. If the points t lie on a straight line, and the planes τ form a pencil, we can obviously find four functions h, k, H, K such that

$$\mu = hA + kB, \qquad \lambda = HA + KB;$$

and these functions *do not depend* upon c. We will say that λ, μ are linear combinations of A, B; thence also $\lambda + \mu, \lambda - \mu$ and consequently

$$A_u - A\,\frac{B_u + A\beta}{B} = B\left\{\left(\frac{A}{B}\right)_u - \beta\left(\frac{A}{B}\right)^2\right\},$$

$$-B_v + B\,\frac{A_v + B\gamma}{A} = \frac{B^2}{A}\left\{\left(\frac{A}{B}\right)_v + \gamma\right\}$$

are such linear combinations. We can also find four functions n, m, ρ, σ of u, v [which do not depend on c] such that:

$$\left(\frac{A}{B}\right)_u = \beta\left(\frac{A}{B}\right)^2 + n\frac{A}{B} + m; \qquad \left(\frac{A}{B}\right)_v = \rho\left(\frac{A}{B}\right)^2 + \sigma\frac{A}{B} - \gamma. \tag{22}$$

Therefore the ratio $A:B$ satisfies a system of Riccati equations. The points t generate a range α on the plane ξ, the planes τ generate a pencil δ of planes pass-

ing through x. As the point x moves on S_x, we obtain ∞^2 ranges α and ∞^2 pencils δ. *All these ranges and all these pencils are projective to one another.*

In the case of Bianchi's permutability theorem, this theorem is evident because the function R of §4 is given by $R = \int dR + c$, and depends only on one additive constant c.

We will say that ∞^1 congruences generate a *pencil* if they have the same first focal surface Σ_x, if homologous points t of the second focal surface Σ_t lie on a straight line, and corresponding tangent planes τ belong to a pencil. We have demonstrated that $r = A:B$ satisfy in this case a Riccati system. If R is a particular solution of this system, the most general solution will be given by

$$r = R\left(1 + \frac{q}{c + s}\right),$$

where q, s are functions of u, v and c is the arbitrary constant. Therefore

$$r_u = -\frac{s_v}{qR}r^2 + \left(\frac{R_u}{R} + \frac{q_u}{q} + 2\frac{s_u}{q}\right)r - R\frac{(s+q)_u}{q},$$

$$r_v = -\frac{s_u}{qR}r^2 + \left(\frac{R_v}{R} + \frac{q_v + 2s_v}{q}\right)r - R\frac{(s+q)_v}{q}.$$

By comparison with (22) we find:

$$-\frac{s_u}{qR} = \beta, \qquad \frac{R(s+q)_v}{q} = \gamma,$$

or

$$\text{(23)} \qquad \begin{cases} s_u = -q\beta R \\ s_v = -q_v + q\gamma R \end{cases}$$

whence

$$\text{(23')} \qquad \left(\frac{q\gamma}{R}\right)_u + (q\beta R)_v = q_{vv}, \qquad s = \int\left[-q\beta R\,du + \left(\frac{q\gamma}{R} - q_v\right)dv\right].$$

For each solution q of this equation we find

$$\text{(24)} \qquad \begin{aligned} r_u &= \beta r^2 + \left(\frac{R_u}{R} + \frac{q_u}{q} - 2\beta R\right)r - R\left(\frac{q_u}{q} - \beta R\right) \\ r_v &= \left(\frac{q_v}{qR} - \frac{\gamma}{R^2}\right)r^2 + \left(\frac{R_v}{R} - \frac{q_v}{q} + 2\frac{\gamma}{R}\right)r - \gamma. \end{aligned}$$

Now let us calculate μ. We find

$$\mu = -A\theta_u - B\left(\frac{A}{B}\right)_u + \frac{A^2}{B}\beta + \left[-B\theta_v + \frac{B^2}{A}\left(\frac{A}{B}\right)_v + \frac{B^2}{A}\gamma\right]$$

$$= -A\theta_u + B(-r_u + \beta r^2) + \left[-B\theta_v + \frac{A}{r^2}(r_v + \gamma)\right]$$

$$= -A\theta_u + B\left\{-\frac{R_u}{R} - \frac{q_u}{q} + 2\beta R\right)r + R\frac{q_u}{q} - \beta R^2\right\}$$

$$+ \left[-B\theta_v + A\left\{\left(\frac{R_v}{R} - \frac{q_v}{q} + 2\frac{\gamma}{R}\right)\frac{1}{r} + \frac{q_v}{qR} - \frac{\gamma}{R^2}\right\}\right]$$

or

$$\mu = A\left(-\theta_u - \frac{R_u}{R} - \frac{q_u}{q} + 2\beta R\right) + B\left(R\frac{q_u}{q} - \beta R^2\right)$$
$$+ A\left(\frac{q_v}{qR} - \frac{\gamma}{R^2}\right) + B\left(-\theta_v + \frac{R_v}{R} - \frac{q_v}{q} + 2\frac{\gamma}{R}\right) \tag{25}$$

For λ we get an analogous value, which we can deduce from the preceding by changing the sign of the two last members.

(25) *is the equation of the straight line composed of the points*

$$\mu x + 2(Ax_u + 2Bx_v)$$

whose local coordinates are μ, A, B.

The analogous equation for λ can be considered as the equation of the pencil of the planes τ.

We transform now equations (23). Let R and R_1 be two solutions of the preceding Riccati system; and, for instance, let us suppose that R $[R_1]$ corresponds to the value $c = \infty$ (to the value $c = 0$) of the arbitrary parameter c. We put $R = A:B$, $R_1 = A_1:B_1$, and choose the arbitrary common factors of A, B and the arbitrary factor of A_1 and B_1 so that

$$r = \frac{Ac + A_1}{Bc + B_1} = \frac{A}{B}\left\{1 + \left(\frac{A_1}{A} - \frac{B_1}{B}\right)\frac{1}{\frac{B_1}{B} + c}\right\}$$

shall be the most general solution of the Riccati system under consideration. In (23) we must therefore write

$$q = \frac{A_1}{A} - \frac{B_1}{B}, \qquad s = \frac{B_1}{B}.$$

Equations (23) are transformed into the following:

$$\frac{\frac{\partial A_1}{\partial v} + \gamma B_1}{A_1} = \frac{\frac{\partial A}{\partial v} + \gamma B}{A},$$
$$\frac{\frac{\partial B_1}{\partial u} + \beta A_1}{B_1} = \frac{\frac{\partial B}{\partial u} + \beta A}{B}. \tag{26}$$

If the congruence determined by $R = A:B$ is W, it is known that we can suppose that the second members of (26) are equal to zero. In this case also the

left-hand members will be equal to zero; and therefore also the congruence determined by $R_1 = A_1:B_1$ or by $r = (A_1 + cA):(B_1 + cB)$ is a W-congruence. Therefore

If one congruence of a pencil is W, all the congruences of the pencil are W. Two W-congruences always determine a pencil of congruences. The determination of a pencil of congruences, if a congruence of the pencil is given, depends on the system (26) *of equations; and these equations are quite analogous to the equations upon which the theory of W congruences depends.*

6. We can also demonstrate the preceding results in another way.

I. Let K be a congruence; the focal points x and y generate the two focal sheets Σ_x, Σ_y; let ξ and η be the planes tangent to Σ_x and Σ_y at two corresponding points x and y. The planes η are identical with the osculating planes of the edges of regression of a system of developables of K; and these edges lie on Σ_x. If a point x moves along such an edge, we have therefore

$$\sum \eta x = \sum \eta\, dx = \sum \eta\, d^2x = 0$$

and consequently also

$$\sum x\, d\eta = \sum dx\, d\eta = 0.$$

Therefore one of the two systems of developables of K satisfies the equations

$$\sum \eta\, dx = 0, \qquad \sum d\eta\, dx = 0.$$

It follows that $\sum d\eta\, dx$ is divisible by $\sum \eta\, dx$, and we put

$$\Omega(x, y) = \frac{\sum dx\, d\eta}{\sum \eta\, dx} = \frac{\sum dx\, d\eta}{-\sum x\, d\eta}.$$

(From $\sum \eta x = 0$, it follows that $\sum \eta\, dx = -\sum x\, d\eta$.) In this equation Ω is a Pfaffian, or a differential which can be exact or not exact. By interchanging the two focal sheets we obtain another Pfaffian

$$\Omega(y, x) = \frac{\sum dy\, d\xi}{\sum \xi\, dy} = \frac{\sum dy\, d\xi}{-\sum y\, d\xi}.$$

By calculation (loc. cit. p. 248) we can verify these theorems, and by a suitable choice of the factors of proportionality of x, y, ξ, η we find that

$$\Omega(y, x) = -\Omega(x, y)$$

and that Ω is *the same* differential $\Omega = \Omega(z, x)$ of (19′), §4. *We get also that, in interchanging the focal sheets, the Pfaffian* Ω *changes only its sign.* If we change the factors of proportionality and substitute ρx, $\sigma y (\rho\sigma \neq 0)$ to x, y, Ω is transformed into

$$\frac{\sum (\rho\, dx + x\, d\rho)(\sigma\, d\eta + \eta\, d\sigma)}{\sigma \sum \eta(x\, d\rho + \rho\, dx)} = \frac{\sum dx\, d\eta}{\sum \eta\, dx} + \left(\frac{d\sigma}{\sigma} - \frac{d\rho}{\rho}\right).$$

In other words, by changing the factors of proportionality, *we can add to* Ω *any exact differential.* The last two theorems give all the possible indeterminations of Ω. We know (l.c.) that K is a W-congruence *only if* Ω *is an exact differential*; and in this case (by a suitable choice of ρ, σ) we can make $\Omega = 0$.

II. Let us now suppose that Σ_z, Σ_y are two surfaces whose points z, y are in one-to-one reciprocal correspondence, and that they determine *a pencil* of ∞^1 surfaces Σ_t; every point t of Σ_t lies on the straight line joining the corresponding points z, y; and the plane τ tangent to Σ_t at t belongs to the pencil determined by the planes ζ, η tangent at z, y to Σ_z, Σ_y.

We can also write

$$t = z + \rho y, \qquad \tau = \zeta + \sigma\eta$$

and the equation $\sum \tau t = \sum \tau\, dt = 0$ become:

$$\rho \sum y\zeta + \sigma \sum \eta z = 0$$

$$d\rho \sum \zeta\, dy + \sigma \sum \eta\, dz = 0.$$

It follows that

$$\frac{d\rho}{\rho} = \frac{\sum \eta\, dz}{\sum \eta z} - \frac{\sum \zeta\, dy}{\sum \zeta y}.$$

The right-hand member must be an exact differential; and, if R is a possible value of ρ, the other values of ρ are cR, where c is an arbitrary constant. And if we substitute Ry for y, we get

$$t = z + cy; \qquad \left[\frac{\sum \eta\, dz}{\sum \eta z} = \frac{\sum \zeta\, dy}{\sum \zeta y}\right].$$

III. Let us suppose that every surface Σ_t is a congruence transform of a fixed surface Σ_x. The lines (x, t) generate a congruence $K(x, t)$ whose focal sheets are Σ_x, Σ_t; and, as Σ_t varies, this congruence generates a *pencil* of congruences. If ξ is the plane tangent at x to Σ_x, and if u, v are curvilinear coordinates on our surfaces, we must have

$$\Omega(z, x) = \frac{\sum d\xi\, dz}{\sum \xi\, dz} = r\, du + s\, dv; \qquad \Omega(y, x) = \frac{\sum d\xi\, dy}{\sum \xi\, dy} = m\, du + n\, dv;$$

$$\Omega(t, x) = \frac{\sum d\xi\, dt}{\sum \xi\, dt} = p\, du + q\, dv,$$

where r, s, m, n are functions *only* of u, v, and p, q are functions of u, v, c. By substituting for t its value $z + cy$, the last equation becomes:

$$\sum d\xi\, dz + c \sum d\xi\, d\eta = (p\, du + q\, dv)\left(\sum \xi\, dz + c \sum \xi\, dy\right)$$

or

$$(r\, du + s\, dv) \sum \xi\, dz + c(m\, du + n\, dv) \sum \xi\, dy$$
$$= (p\, du + q\, dv)\left[\sum \xi\, dz + c \sum \xi\, dy\right].$$

But $\sum \xi\, dy$ and $\sum \xi\, dz$ are not proportional (because the surfaces Σ_y and Σ_z are not identical). And therefore

$$(r\, du + s\, dv) - (p\, du + q\, dv) = cR \sum \xi\, dy$$

$$(p\, du + q\, dv) - (m\, du + n\, dv) = R \sum \xi\, dz,$$

where R is a function of u, v, c.

By summing we get

$$(r\, du + s\, dv) - (m\, du + n\, dv) = cR \sum \xi\, dy + R \sum \xi\, dz.$$

By derivation with respect to c, we obtain

$$\frac{\partial(cR)}{\partial c} \sum \xi\, dy + \frac{\partial R}{\partial c} \sum \xi\, dz = 0;$$

therefore R and Rc do not depend upon c; and consequently $R = 0$. We get also

$$r\, du + s\, dv = m\, du + n\, dv = p\, du + q\, dv,$$

that is

$$\Omega(z, x) = \Omega(y, x) = \Omega(t, x),$$

and these equations give the theorem on the pencils of congruences we have demonstrated above (§5).

IV. Let us suppose now that the surfaces Σ_y, Σ_z, Σ_t are also congruence transforms of another surface Σ_w. From the theorem of §4 we deduce:

$$\Omega(t, w) + \Omega(t, x) = 0.$$

Also, if only one of the differentials $\Omega(t, x)$, $\Omega(t, w)$ is exact, all those differentials [which are all equal to $\Omega(z, w)$ or $\Omega(z, x)$] are exact. This is precisely the first theorem of §1.

V. Let us suppose now that the surfaces Σ_y, Σ_z, Σ_t are congruence transforms not only of Σ_x, Σ_w, but also of the surfaces Σ_j of the pencil determined by Σ_x, Σ_w. We have in such a case

$$0 = \Omega(y, x) + \Omega(y, j) = \Omega(y, j) + \Omega(y, w) = \Omega(y, w) + \Omega(y, x).$$

Therefore all the Ω are equal to zero,[10] and all the congruences are W congruences. This is precisely the second theorem of §1.

In a preceding paper of mine, *On a property of W-congruences* (these Annals, vol. 41, no. 2 (April 1940) pp. 356–364) the following corrections are to be made on page 360. In line 3 replace $A_u + A\theta_u$ by $Az_u + Bz_v$; also in line 3 replace $B_v + B\theta_v$ by $Az_u + Bz_v$; and in line 6 replace x by y.

INSTITUTE FOR ADVANCED STUDY.

[10] We could also say that the preceding sums are exact differentials, and that consequently also the Ω are exact differentials.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

ON CONTINUOUS MAPPING

BY NELSON DUNFORD

(Received June 17, 1939)

1. Introduction

In this paper we are concerned with conditions which are sufficient to insure the existence of an interior point in the map $f(G)$ of a non-void open set G under a 1-1 continuous function f. The functions f actually considered in certain of our results form an extension of the notion of the class of 1-1 maps. The term *distributive* has been used to describe this more general type of function for the reason that every homomorphic[1] map between metric groups with a one-sided invariant metric is an instance of such a function. In the case of continuous homomorphic maps on groups or continuous additive operations on a linear space it is not difficult to show that if the map of every open set contains an interior point then the mapping must be interior, i.e., take open sets into open sets; and this interior property of linear maps is of prime importance in the theory of linear operations. In the general case we are considering here one cannot hope for so much, for it is easy to give examples of 1-1 maps between complete separable and locally compact metric spaces which transform non-void open sets into sets with an interior point and yet are not interior maps. There are results of the type we are considering which may be readily obtained from certain very restrictive assumptions. For example if the range space satisfies the first Hausdorff axiom of countability and the domain space is regular and locally compact it is easily seen that any continuous map which transforms open sets into sets which are not non-dense must actually take every non-void open set into the sum of an open set and a non-dense set. This and other similar results which may be obtained from a compactness assumption have been omitted since they are comparatively trivial. In no place do we assume that the range space satisfies the first axiom of countability and in no place do we assume local compactness on either domain or range space.

The paper divides itself into three parts. In the first (§§3, 4) we are concerned with placing certain results which are well known in the theory of metric spaces or in the theory of the neighborhood spaces of Hausdorff on a broader basis. These results center around the notions of category, the condition of Baire, and the operation A of Souslin. Since the relations between these notions are of considerable importance in many phases of topology we develop them here for a general type of neighborhood space in which the product of two open sets need not be open. Rather than define the closure operation by means of the neigh-

[1] I.e., $f(x + y) = f(x) + f(y)$ if the groups are written additively.

borhoods we introduce it postulationally but later in certain of the results in §5, assume that it is the ordinary closure determined by the neighborhoods.

The second part of the paper concerns itself with continuous functions f. In order to get an idea of what sort of conditions might be imposed in order that the map $f(G)$ of an open set contain an interior point we examine the known[2] instance of a continuous homomorphic map of one complete metric separable group upon the whole of another such group. Here, as shown by Banach, the map must be interior. In this situation we see immediately (in view of the Baire category theorem) that every non-void open set G transforms into a set $f(G)$ which is not non-dense. This condition, which is obviously a necessary one, is quite a long way from the desired conclusion, i.e., $f(G)$ contains an interior point. In this case considered by Banach we can obtain considerably more as a starting point, namely that $f(G)$ (G open and non-void) is a Baire set of the second category. This is because the map, being between complete separable spaces, takes Borel sets into analytic sets and such sets satisfy the condition of Baire. That $f(G)$ is actually a second category Baire set follows immediately from the Baire category theorem and the additive nature of f. The fact that $f(G)$ is a second category Baire set seems to be considerably nearer the desired conclusion than the statement that $f(G)$ is not non-dense. These two observations give us the clue used in §5. In the first two theorems of §5 we have assumed something slightly less restrictive than the assertion that $f(G)$ is a second category Baire set. Under such an assumption on f very little need be assumed about the domain and range spaces. In Theorem 5.5 we have assumed only that $f(G)$ is not non-dense. In this situation we have had to restrict the spaces concerned considerably. The range space is assumed to be a Hausdorff space and the domain space locally complete and locally separable. It might be mentioned here (and also used as an example showing the necessity of the 1–1 assumption in Theorem 5.1) that a continuous function $y = f(x)$ may be defined on a separable metric space X and have for its range the whole of the line segment $0 \leqq y \leqq 1$ (with its usual topology) and take every open set into a Baire set of second category and still have the interior of $f(S)$ empty for every sufficiently small sphere S. To construct the space X we may take two unit line segments in the plane each directly above the segment $(0, 1)$. On one segment remove all rational points and on the other remove all points of another denumerable dense set which contains no rationals. The metric in X is the usual distance in the plane and the function $y = f(x)$ is a projection of x onto the segment $(0, 1)$. In the last part of the paper, §6, an application is made.

Finally we wish to point out that the method of proof used in Theorem 3.1 is due to Sierpinski.[3]

---

[2] S. Banach, *Über metrische Gruppen*, Studia Math., vol. 3(1931), pp. 101–113, in particular Theorem 8.

[3] W. Sierpinski, *General Topology*, University of Toronto press (1934), p. 194.

2. Notation, axioms, and definitions

Throughout the paper we shall use the following notation: $X = (x)$ is a set of points, $\mathfrak{P}(X)$ is the power set of X, i.e., the family of all subsets of X, $\mathfrak{N} \subset \mathfrak{P}(X)$ is a family of sets called neighborhoods, $\bar{E}$ called closure of E, is a function on $\mathfrak{P}(X)$ to $\mathfrak{P}(X)$. The symbol N with or without subscripts or superscripts will always denote a neighborhood. A set whose closure contains no N is said to be non-dense. Later we shall restrict the family $\mathfrak{N}$ and the closure operation $\bar{E}$ according to a practise in common use but for the present the following less restrictive axioms are sufficient.

A(o) $\mathfrak{N} \subset \mathfrak{P}(X)$
(i) *$N \in \mathfrak{N}$ implies $N \neq 0$, i.e., not void.*
(ii) *$N_1N_2 \neq 0$ implies the existence of an $N_3 \subset N_1N_2$.*
B(i) *Any subset of a non-dense set is non-dense.*
(ii) *$N \subset \bar{E}$ implies $N \subset \overline{NE}$.*
(iii) *The void set is non-dense.*

It should be noted that the N_3 in A(ii) is not required to contain a specified point of N_1N_2. In fact, until §5, we shall find it convenient not to mention the points x of X but deal entirely with the sets $E \in \mathfrak{P}(X)$. For this reason a few of the definitions to follow may sound a bit unusual.

The first small greek letters α, β, γ, δ, ϵ, will be used for ordinals and the last ones ϕ, χ, ψ, ω, when written without primes, for cardinals. The axiom of choice will be used throughout the paper and the symbol ϕ' will always mean the smallest ordinal corresponding to the cardinal ϕ.

The family of sets E in $\mathfrak{P}(X)$ which can be expressed as the sum of ϕ non-dense sets will be denoted by I_ϕ, and we define $II_\phi = \mathfrak{P}(X) - I_\phi$. A set E is said to be *locally I_ϕ* if for every non-void subset E' of E there is an N with $NE' \neq 0$, $NE \in I_\phi$.

3. Category theorems and the class $\mathfrak{B}_\phi$

THEOREM 3.1: *If E is locally I_ϕ then $E \in I_\phi$.*

Well-order all N having the property that $NE \in I_\phi$;

$$N_1, N_2, \cdots, N_\alpha, \cdots \qquad \alpha < \beta. \tag{3.1}$$

Define

$$M_1 = N_1E, \qquad M_\alpha = N_\alpha E - \sum_{\delta<\alpha} N_\delta. \tag{3.2}$$

It follows from B(i) that subsets of I_ϕ sets are also I_ϕ sets and hence we may write

$$M_\alpha = \sum_{\gamma<\phi'} M_\alpha^\gamma \qquad \text{where } M_\alpha^\gamma \text{ is non-dense.}$$

Now define

$$E^\gamma = \sum_{\alpha<\beta} M_\alpha^\gamma.$$

It will first be shown that

$$E = \sum_{\gamma<\phi'} E^{\gamma}. \tag{3.3}$$

Obviously $E \supset \sum_{\gamma<\phi'} E^{\gamma}$. Now if $E_1 = E - \sum_{\gamma<\phi'} E^{\gamma} \neq 0$ there is an N with $NE_1 \neq 0$, $NE \in I_{\phi}$. Thus let N_{ϵ} be the first in the series (3.1) such that $N_{\epsilon}E_1 \neq 0$. Then

$$M_{\epsilon} = N_{\epsilon}E - \sum_{\delta<\epsilon} N_{\delta} \supset N_{\epsilon}E_1 \neq 0,$$

so that

$$0 \neq N_{\epsilon}E_1 = N_{\epsilon}E_1M_{\epsilon} = \sum_{\gamma<\phi'} N_{\epsilon}E_1M_{\epsilon}^{\gamma}.$$

Hence there is a $\gamma < \phi'$ such that $N_{\epsilon}E_1M_{\epsilon}^{\gamma} \neq 0$, a fortiori $N_{\epsilon}E_1E^{\gamma} \neq 0$, $N_{\epsilon}E_1 \sum_{\gamma<\phi'} E^{\gamma} \neq 0$, which is a contradiction to the definition of E_1. This establishes equation (3.3). It remains therefore to be shown that the set $E^{\gamma}(\gamma < \phi')$ is non-dense. Suppose on the contrary that for some $\gamma < \phi'$, $\bar{E}^{\gamma} \supset N$. By B(ii) then $N \subset \overline{NE^{\gamma}}$ which shows (using B(i) and the fact that the void set is non-dense) that there is some α for which $NN_{\alpha} \neq 0$. Let N_{ϵ} be the first in the series (3.1) such that $NN_{\epsilon} \neq 0$. There is by A(ii) an $N_* \subset NN_{\epsilon}$ and we have $N_*N_{\alpha} = 0$ $(\alpha < \epsilon)$, a fortiori $N_*M_{\alpha} = 0$ $(\alpha < \epsilon)$. But from (3.2) it follows that $N_{\epsilon}M_{\alpha} = 0$ for $\alpha > \epsilon$ and thus $N_*M_{\alpha} = 0$ for $\alpha > \epsilon$. Thus

$$N_*M_{\alpha}^{\gamma} = 0, \qquad \alpha \neq \epsilon,$$

and

$$N_*E^{\gamma} = N_* \sum_{\alpha<\beta} M_{\alpha}^{\gamma} = N_*M_{\epsilon}^{\gamma} \subset M_{\epsilon}^{\gamma}. \tag{3.4}$$

Now

$$N_* \subset NN_{\epsilon} \subset N \subset \bar{E}^{\gamma}$$

and therefore $N_* \subset \overline{N_*E^{\gamma}}$, a contradiction to B(i), (3.4), and the fact that M_{ϵ}^{γ} is a non-dense set. This completes the proof of Theorem 3.1.

For an arbitrary set E we define E_{ϕ}^* as the sum of all subsets e of E for which there exists an $N \supset e$ such that $NE \in I_{\phi}$.

THEOREM 3.2: *For an arbitrary set E the set E_{ϕ}^* is a I_{ϕ} set.*

In view of Theorem 3.1 it is sufficient to show that E_{ϕ}^* is locally I_{ϕ}. Let $0 \neq e' \subset E_{\phi}^*$. There exists an e (one of the summands in $E_{\phi}^* = \sum e$) such that $ee' \neq 0$. Thus some $N \supset e$ and $NE \in I_{\phi}$. So $Ne' \neq 0$ and NE_{ϕ}^*, being a subset of NE, must be a I_{ϕ} set.

A set E is said to be *locally* II_{ϕ} on a set e if for every N with $Ne \neq 0$ we have $NE \in II_{\phi}$.

THEOREM 3.3: *If ϕ is transfinite and if $E \in II_{\phi}$ there is an $N_0 \in \mathfrak{N}$ such that E is locally II_{ϕ} on N_0. Furthermore if one assumes that the sum of two non-dense sets is non-dense then this is true also for ϕ finite.*

For some N_0 we have

(3.5) $$\overline{E - E^*_\phi} \supset N_0 ,$$

for if not, $E - E^*_\phi$ is non-dense and by Theorem 3.2 E is either non-dense, which is impossible, or else $E = E - E^*_\phi + E^*_\phi$ is the sum of $\phi + 1$ non-dense sets. If ϕ is transfinite then $\phi + 1 = \phi$ and $E \epsilon I_\phi$, a contradiction. Also if ϕ is finite and the sum of two (and thus any finite number) of non-dense sets is non-dense then E is non-dense, which again is impossible. Thus we may assume the existence of an N_0 satisfying (3.5). Now suppose $NN_0 \neq 0$ and pick $N_1 \subset NN_0$; then

$$0 \neq N_1 \subset NN_0 \subset N_0 \subset \overline{E - E^*_\phi},$$

and so

$$N_1 \subset \overline{N_1(E - E^*_\phi)}.$$

By B(iii) then

$$0 \neq N_1(E - E^*_\phi) \subset N(E - E^*_\phi).$$

Since $N(E - E^*_\phi)E^*_\phi = 0$ there is no subset e of $N(E - E^*_\phi)$ having the property that there exists an $N' \supset e$ with $N'E \epsilon I_\phi$. In particular, since $N \supset N(E - E^*_\phi)$, we have $NE \epsilon II_\phi$ which shows that E is locally II_ϕ on N_0.

Any set which is either void or the sum of neighborhoods will be called *open* and the complements of open sets will be called *closed*.[4] The letter $\mathfrak{G}$, $[\mathfrak{F}]$ will be used for the family of open [closed] sets. The symbol $\mathfrak{B}_\phi$ will be used for the family of sets E having the property that $N - E \epsilon I_\phi$ whenever E is locally II_ϕ on N. From this point on we shall assume

C *Either ϕ is transfinite or the sum of two non-dense sets is non-dense.*

THEOREM 3.4: *The family $\mathfrak{B}_\phi$ has the following properties.*

(i) *If E_α $(\alpha < \phi')$ is a $\mathfrak{B}_\phi$ set so is $\sum_{\alpha<\phi'} E_\alpha$.*

(ii) *The complement of a $\mathfrak{B}_\phi$ set is a $\mathfrak{B}_\phi$ set and the product and difference of two $\mathfrak{B}_\phi$ sets are $\mathfrak{B}_\phi$ sets.*

(iii) *Every open and every closed set is a $\mathfrak{B}_\phi$ set.*

(iv) *For an arbitrary set E there is a set $B \epsilon \mathfrak{B}_\phi$ covering E and such that if B' is any $\mathfrak{B}_\phi$ set covering E then $B - B' \epsilon I_\phi$.*

To prove (i) suppose

(a) $$\sum_{\alpha<\phi'} E_\alpha \text{ is locally } \mathrm{II}_\phi \text{ on } N.$$

We have to show that $N - \sum_{\alpha<\phi'} E_\alpha \epsilon I_\phi$. If this were not the case there would be (using C and Theorem 3.3) an N_0 such that

(b) $$N - \sum_{\alpha<\phi'} E_\alpha \text{ is locally } II_\phi \text{ on } N_0 .$$

[4] Note that the closure $\bar{E}$ of a set E is not necessarily a closed set.

From (b) and C, $NN_0 \neq 0$ and from (a) $N_0 \sum_{\alpha<\phi'} E_\alpha \epsilon II_\phi$. Thus there is an α_0 such that $N_0 E_{\alpha_0} \epsilon II_\phi$ and (Theorem 3.3) an N_0' such that $N_0 E_{\alpha_0}$ is locally II_ϕ on N_0'. Then, *a fortiori*,

(c) $$E_{\alpha_0} \text{ is locally } II_\phi \text{ on } N_0'.$$

Since $E_{\alpha_0} \epsilon \mathfrak{B}_\phi$ we have from (c) that $N_0' - E_{\alpha_0} \epsilon I_\phi$. But $N_0' - E_{\alpha_0} \supset N_0' - \sum_{\alpha<\phi'} E_\alpha$ and thus

$$N_0' - \sum_{\alpha<\phi'} E_\alpha \epsilon I_\phi,$$

which, in view of the fact that $N_0'N_0 \neq 0$, is a contradiction to (b). This completes the proof of (i).

To prove (ii) let N be such that

(d) $$X - E \text{ is locally } II_\phi \text{ on } N.$$

We wish to conclude that the set $NE = N - (X - E)$ is a I_ϕ set. If $NE \epsilon II_\phi$ there is an N_0 such that NE is locally II_ϕ on N_0 and thus E is locally II_ϕ on N_0. Since $E \epsilon \mathfrak{B}_\phi$, $N_0 - E \epsilon I_\phi$ which is a contradiction to (d) and the fact that $NN_0 \neq 0$. The rest of (ii) follows immediately from (i) and the formulas

$$ED = X - [(X - E) + (X - D)],$$
$$E - D = E(X - D).$$

To prove (iii) let F be a closed set so that $F = X - G$ where G is an open set and let N be such that

(e) F is locally II_ϕ on N.

Then if $N - F = NG \epsilon II_\phi$, it is not empty and there is an $N' \subset NG$; thus by (e) the void set $FN' \epsilon II_\phi$ which is impossible. Thus closed sets are $\mathfrak{B}_\phi$ sets and by (ii) open sets are also.

To prove (iv) let E be an arbitrary set and let G be the sum of all N such that $NE \epsilon I_\phi$. Place

$$F = X - G, \qquad B = F + GE,$$

then to see that the covering B of E is a $\mathfrak{B}_\phi$ set it is sufficient in view of (i), (iii), and the fact that $I_\phi \subset \mathfrak{B}_\phi$, to show that $GE \epsilon I_\phi$; and in view of B(iii) and Theorem 3.1 it is sufficient for this to show that GE is locally I_ϕ. This last fact is an immediate consequence of the definition of G. Now let B' be an arbitrary $\mathfrak{B}_\phi$ covering of E, G' the sum of all N such that $NB' \epsilon I_\phi$, and $F' = X - G'$. Thus $G' \subset G$, $F' \supset F$ and

$$B - B' = F + GE - B' = F - B' \subset F' - B'.$$

To complete the proof of (iv) it is sufficient then to show that $F' - B' \epsilon I_\phi$. Suppose on the contrary that $F' - B' \epsilon II_\phi$ and thus $F' \epsilon II_\phi$. Let F' be locally II_ϕ on N; then if $N \not\subset F'$ there is an $N' \subset NG'$, and thus $0 = N'F' \epsilon II_\phi$,

which is a contradiction to B(iii). Thus $F' \supset N$. It follows readily from the definition of G' that B' is locally II_ϕ on F' and thus since $B' \in \mathfrak{B}_\phi$ we have

(f) $$N(F' - B') \in I_\phi \text{ for every } N \subset F'.$$

Now since $F' - B'$ was assumed to be II_ϕ there is an N_0 such that it is locally II_ϕ on N_0, and thus in view of (f), $N_0G' \neq 0$. It follows that there is an $N' \subset N_0G'$ and since $F' - B'$ is locally II_ϕ on N_0 we have the null set $0 = N'(F' - B') \in II_\phi$, a contradiction to B(iii). This completes the proof of Theorem 3.4.

4. The operation A_{Φ_0}

Let Φ be the set of all denumerable sequences $\mathfrak{a} = \{\alpha_n\}$ of ordinals $< \phi'$, i.e., $n = 1, 2, \cdots$ implies $\alpha_n < \phi'$. For two different sequences $\mathfrak{a} = \{\alpha_n\}$, $\mathfrak{a}' = \{\alpha_n'\}$ in Φ let the distance between $\mathfrak{a}$ and $\mathfrak{a}'$ be defined by the formula

$$(\mathfrak{a}, \mathfrak{a}') = 1/m[\mathfrak{a}, \mathfrak{a}'],$$

where $m[\mathfrak{a}, \mathfrak{a}']$ is the first integer m for which $\alpha_m \neq \alpha_m'$. This metric of Baire makes the space Φ, as is well known (and readily verified), a complete metric space. Suppose now that for every $\mathfrak{a} = \{\alpha_n\}$ in a certain subset Φ_0 of Φ and every integer n there is a set $K_{\alpha_1 \cdots \alpha_n}$ which, as the notation indicates, depends merely upon the finite subsequence $\alpha_1, \alpha_2, \cdots, \alpha_n$ of $\mathfrak{a}$ and is thus not an arbitrary function of $\mathfrak{a}$ and n. Such a function $K = K_{\alpha_1 \cdots \alpha_n}$ is called a kernel over Φ_0. For every kernel K over Φ_0 the operation

$$A_{\Phi_0}[K] = \sum_{\mathfrak{a} \in \Phi_0} \prod_{n=1}^{\infty} K_{\alpha_1 \cdots \alpha_n}$$

will then be defined. If $\mathfrak{M} \subset \mathfrak{P}(X)$, by $A_{\Phi_0}[\mathfrak{M}]$ will be meant the family of all sets $A_{\Phi_0}[K]$ where $K = K_{\alpha_1 \cdots \alpha_n}$ is a kernel over Φ_0 with values in $\mathfrak{M}$.

THEOREM 4.1: *If $\phi \geqq \aleph_0$ and Φ_0 is the sum of at most ϕ closed sets in Φ then $\mathfrak{B}_\phi$ is an invariant of the operation A_{Φ_0}, i.e.,*

$$A_{\Phi_0}[\mathfrak{B}_\phi] = \mathfrak{B}_\phi .$$

While it follows from the theorem that $\mathfrak{B}_\phi$ is an invariant of A_{Φ_0} for every open set Φ_0 and every F_σ set Φ_0 in Φ we do not know if this is true if Φ_0 is, for example, an arbitrary Borel set in Φ.

Before proving the theorem we would like to point out that if Φ_0 is closed in Φ and $\mathfrak{a} = \{\alpha_n\} \in \Phi - \Phi_0$ then there is an $m = m(\mathfrak{a}, \Phi_0)$ such that for every $\mathfrak{b} = \{\beta_n\} \in \Phi_0$ one of the inequalities

$$\beta_n \neq \alpha_n \qquad n = 1, 2, \cdots, m$$

holds. This shows that a kernel K over a closed Φ_0 may be extended so as to be defined over all of Φ and satisfying the following conditions

(i) The extended kernel K' takes on no values that K does not assume with perhaps the exception of the void set.

(ii) $$A_{\Phi_0}[K] = A_{\Phi}[K'].$$

To define $K'_{\alpha_1 \cdots \alpha_n}$ we place $K'_{\alpha_1 \cdots \alpha_n} = 0$ if there is no sequence in Φ_0 starting with $\alpha_1, \alpha_2, \cdots, \alpha_n$ otherwise place $K'_{\alpha_1 \cdots \alpha_n} = K_{\alpha_1 \cdots \alpha_n}$. Since Φ_0 is closed we have for every $\mathfrak{a} \in \Phi - \Phi_0$ the product $\prod_{n=1}^{\infty} K'_{\alpha_1 \cdots \alpha_n}$ void, and thus (ii) is true. As the reader will see presently the theorem when Φ_0 is the sum of ϕ closed sets is an immediate corollary of the case where Φ_0 is closed. Thus by the above remarks it would be sufficient to prove the theorem for the single case $\Phi_0 = \Phi$. This can be proved by very slight modification of methods already known.[5] Rather than use this argument we prefer to give a complete proof here for an arbitrary Φ_0 since this attack will bring out more clearly the one place in the proof where an assumption on Φ_0 is needed and may therefore throw some light on the general case of $\Phi_0 =$ an arbitrary Borel set.

Proceeding now to the proof of the theorem we let K be a kernel over Φ_0 with values in $\mathfrak{B}_\phi$, where Φ_0 will for the time being, be subject to no restrictions. It may be supposed that for $\mathfrak{a} = \{\alpha_n\} \in \Phi_0$ we have

$$K_{\alpha_1 \cdots \alpha_n} \supset K_{\alpha_1 \cdots \alpha_{n+1}} \qquad n = 1, 2, \cdots, \tag{4.1}$$

for if this were not true the kernel

$$K' = K'_{\alpha_1 \cdots \alpha_n} = \prod_{m=1}^{n} K_{\alpha_1 \cdots \alpha_m}$$

has the property of being regular expressed by 4.1, has, by Theorem 3.4 (ii), its values in $\mathfrak{B}_\phi$, and $A_{\Phi_0}[K'] = A_{\Phi_0}[K]$.

By Theorem 3.4 (iv) there is a set $B \in \mathfrak{B}_\phi$ such that

$$A_{\Phi_0}[K] \subset B, \text{ and} \tag{4.2}$$

$$\text{If } B' \in \mathfrak{B}_\phi \text{ and } A_{\Phi_0}[K] \subset B' \text{ then } B - B' \in I_\phi. \tag{4.3}$$

For $\mathfrak{b} = \{\beta_n\} \in \Phi_0$ and $i = 1, 2, \cdots$ let $\Phi_0(\mathfrak{b}, i)$ be the set of $\mathfrak{a} = \{\alpha_n\}$ in Φ for which

$$(\beta_1, \beta_2, \cdots, \beta_i, \alpha_1, \alpha_2, \cdots) \in \Phi_0.$$

Again by Theorem 3.4 (iv) there is for each $\mathfrak{b} = \{\beta_n\} \in \Phi_0$ and each $i = 1, 2, \cdots$ a set $B_{\beta_1 \cdots \beta_i} \in \mathfrak{B}_\phi$ depending only upon the finite sequence $\beta_1, \cdots, \beta_i$ and such that

$$\sum_{\mathfrak{a} \in \Phi_0(\mathfrak{b}, i)} \prod_{n=1}^{\infty} K_{\beta_1 \cdots \beta_i \alpha_1 \cdots \alpha_n} \subset B_{\beta_1 \cdots \beta_i}, \tag{4.4}$$

and

$$\text{if } B' \in \mathfrak{B}_\phi \text{ and } \sum_{\mathfrak{a} \in \Phi_0(\mathfrak{b}, i)} \prod_{n=1}^{\infty} K_{\beta_1 \cdots \beta_i \alpha_1 \cdots \alpha_n} \subset B', \text{ then } B_{\beta_1 \cdots \beta_i} - B' \in I_\phi. \tag{4.5}$$

[5] See, for example, Kuratowski, *Topologie* I, Warsaw 1933; pp. 56–57.

It may be supposed that

$$B_{\beta_1\cdots\beta_i} \subset K_{\beta_1\cdots\beta_i} \quad \text{for} \quad \mathfrak{b} \in \Phi_0, \qquad i = 1, 2, \cdots, \tag{4.6}$$

for in view of 4.1 and the fact that the product of two $\mathfrak{B}_\phi$ sets is a $\mathfrak{B}_\phi$ set, the set $K_{\beta_1\cdots\beta_i}B_{\beta_1\cdots\beta_i}$ satisfies all the conditions imposed on $B_{\beta_1\cdots\beta_i}$. Since by 4.2

$$A_{\Phi_0}[K] = B - (B - A_{\Phi_0}[K])$$

and B is a $\mathfrak{B}_\phi$ set it suffices to show that $B - A_{\Phi_0}[K] \in I_\phi$. Using 4.6

$$B - A_{\Phi_0}[K] = B - \sum_{\mathfrak{b} \in \Phi_0} \prod_{i=1}^{\infty} K_{\beta_1\cdots\beta_i} \subset B - \sum_{\mathfrak{b} \in \Phi_0} \prod_{i=1}^{\infty} B_{\beta_1\cdots\beta_i}. \tag{4.7}$$

For $\mathfrak{b} = \{\beta_n\} \in \Phi_0$, and $i = 1, 2, \cdots$ let $(\mathfrak{b}, i)$ be the set of ordinals $\gamma < \phi'$ for which there is a sequence of the form

$$(\beta_1, \beta_2, \cdots, \beta_i, \gamma, \cdots) \in \Phi_0.$$

For $i = 0$ the symbol $(\mathfrak{b}, i)$ is taken to mean the set of all ordinals $\gamma < \phi'$ for which there exists a sequence $(\gamma, \cdots)$ starting with γ and belonging to Φ_0. With this notation in mind we wish to establish the following inclusion:

$$B - \sum_{\mathfrak{b} \in \Phi_0} \prod_{i=1}^{\infty} B_{\beta_1\cdots\beta_i} \subset \sum_{\mathfrak{b} \in \Phi_0} \sum_{i=0}^{\infty} \Big(B_{\beta_1\cdots\beta_i} - \sum_{\gamma \in (\mathfrak{b}, i)} B_{\beta_1\cdots\beta_i\gamma}\Big), \tag{4.8}$$

where $B_{\beta_1\cdots\beta_i} = B$ for $i = 0$. This inclusion will now be established under the assumption that Φ_0 is a closed set in Φ and the theorem will be proved for this case first.

Suppose x is a point of B but does not belong to the right side of (4.8). Then for every $\mathfrak{b} \in \Phi_0$ we have the following statement holding.

(4.9) If $x \in B_{\beta_1\cdots\beta_i}$ there will be a $\gamma \in (\mathfrak{b}, i)$ such that $x \in B_{\beta_1\cdots\beta_i\gamma}$.

Since it is known that $x \in B$ we get, upon taking $i = 0$ in 4.9, a $\gamma_1 \in (\mathfrak{b}, 0)$ such that $x \in B_{\gamma_1}$. Since $\gamma_1 \in (\mathfrak{b}, 0)$ there is a sequence $\mathfrak{b}_1 = (\gamma_1, \cdots)$ in Φ_0. For $\mathfrak{b} = \mathfrak{b}_1$ we have, upon setting $i = 1$ in (4.9), a γ_2 in $(\mathfrak{b}_1, 1)$ such that $x \in B_{\gamma_1\gamma_2}$. Since γ_2 is in $(\mathfrak{b}_1, 1)$ there is a sequence $\mathfrak{b}_2 = (\gamma_1, \gamma_2, \cdots)$ in Φ_0 and we can repeat the argument and arrive at a sequence

$$\mathfrak{c} = (\gamma_1, \gamma_2, \gamma_3, \cdots) \in \Phi, \qquad x \in B_{\gamma_1\cdots\gamma_n} \quad (n = 1, 2, \cdots)$$

and the sequence $\mathfrak{c}$ has the property that for every integer n there is a point $\mathfrak{a}_n \in \Phi_0$ of the form $\mathfrak{a}_n = (\gamma_1, \gamma_2, \cdots, \gamma_n, \alpha_{n+1}, \alpha_{n+2}, \cdots)$. Since it is assumed that Φ_0 is closed in the Baire metric of Φ we see that the sequence $\mathfrak{c}$ belongs to Φ_0 and that consequently

$$x \in \sum_{\mathfrak{b} \in \Phi_0} \prod_{i=1}^{\infty} B_{\beta_1\cdots\beta_i},$$

which establishes the inclusion (4.8).

Since as $\mathfrak{b}$ ranges over Φ_0 there are at most $\phi^i = \phi$ sets

$$B_{\beta_1 \cdots \beta_i} - \sum_{\gamma \in (\mathfrak{b}, i)} B_{\beta_1 \cdots \beta_i \gamma}, \tag{4.10}$$

there will be at most $\aleph_0 \phi = \phi$ sets in the $\sum \sum$ on the right side of (4.8). Hence to show that $B - A_{\Phi_0}[K] \in I_\phi$ it suffices, in view of (4.7) and (4.8), to show that 4.10, is a I_ϕ set. Now if

$$x \in \sum_{\mathfrak{a} \in \Phi_0(\mathfrak{b}, i)} \prod_{n=1}^{\infty} K_{\beta_1 \cdots \beta_i \alpha_1 \cdots \alpha_n}, \tag{4.11}$$

there is a sequence $\mathfrak{a} = \{\alpha_n\}$ such that

$$\mathfrak{b} = (\beta_1, \beta_2, \cdots, \beta_i, \alpha_1, \alpha_2, \cdots) \in \Phi_0, \tag{4.12}$$

and

$$x \in K_{\beta_1 \cdots \beta_i \alpha_1 \cdots \alpha_n} \qquad n = 1, 2, \cdots. \tag{4.13}$$

Let $\mathfrak{b}$ be the sequence in (4.12), then upon using (4.4) with i replaced by $i + 1$ we have

$$\sum_{\mathfrak{a}' \in \Phi_0(\mathfrak{b}, i+1)} \prod_{n=1}^{\infty} K_{\beta_1 \cdots \beta_i \alpha_1 \alpha_1' \alpha_2' \cdots \alpha_n'} \subset B_{\beta_1 \cdots \beta_i \alpha_1}. \tag{4.14}$$

Using (4.12) and (4.13) we see that $\mathfrak{a}' = (\alpha_2, \alpha_3, \cdots) \in \Phi_0 (\mathfrak{b}, i + 1)$ and thus from (4.11) and (4.14) that $x \in B_{\beta_1 \cdots \beta_i \alpha_1}$, which proves that

$$\sum_{\mathfrak{a} \in \Phi_0(\mathfrak{b}, i)} \prod_{n=1}^{\infty} K_{\beta_1 \cdots \beta_i \alpha_1 \cdots \alpha_n} \subset \sum_{\gamma \in (\mathfrak{b}, i)} B_{\beta_1 \cdots \beta_i \gamma}. \tag{4.15}$$

Since there are at most ϕ terms in the sum $\sum_{\gamma \in (\mathfrak{b}, i)} B_{\beta_1 \cdots \beta_i \gamma}$ and each term is a $\mathfrak{B}_\phi$ set we have by Theorem 3.4 (i) and (4.15) the right to use this sum as B' in (4.5) (or (4.3) in case $i = 0$) and thus conclude that (4.10) is an I_ϕ set. Thus it has been shown that if Φ_0 is closed and the kernel K over Φ_0 has its values in $\mathfrak{B}_\phi$ then $A_{\Phi_0}[K] \in \mathfrak{B}_\phi$ and so for closed Φ_0, $A_{\Phi_0}[\mathfrak{B}_\phi] \subset \mathfrak{B}_\phi$. On the other hand it is obvious that $A_{\Phi_0}[\mathfrak{B}_\phi] \supset \mathfrak{B}_\phi$. This establishes the theorem for closed sets Φ_0. In view of Theorem 3.4 (i) and the formula

$$\sum_{\mathfrak{a} \in \sum_{\gamma < \phi'} \Phi_\gamma} \prod_{n=1}^{\infty} K_{\alpha_1 \cdots \alpha_n} = \sum_{\gamma < \phi'} \sum_{\mathfrak{a} \in \Phi_\gamma} \prod_{n=1}^{\infty} K_{\alpha_1 \cdots \alpha_n}$$

it follows immediately that $A_{\Phi_0}[\mathfrak{B}_\phi] = \mathfrak{B}_\phi$ for any set Φ_0 which is the sum of at most ϕ closed sets.

5. Continuous functions

Throughout this section we shall be dealing with two spaces X, X_1 and continuous mappings $f(e)$ on $\mathfrak{P}(X)$ to $\mathfrak{P}(X_1)$. Our chief interest will lie in the discussion of conditions sufficient to insure that the interior of the image of a

neighborhood is not empty. In the first theorems to follow there is very little assumed explicitly about the spaces X, X_1. What gives the hold on the situation in these theorems is the assumption involving implicitly X, X_1, and f and asserting that for every $N \epsilon \mathfrak{N}$ there is a cardinal number $\phi(N)$, finite or transfinite, such that $f(N) \epsilon \mathfrak{B}_{\phi(N)} II_{\phi(N)}$. In the last part of this section we investigate conditions under which the existence of such a function $\phi(N)$ can be established.

By the *ordinary closure* $\bar{E}$ of a set E we shall mean the sum of all sets e having the property that $Ne \neq 0$ implies $NE \neq 0$. We shall find it necessary at times to restrict the space X or X_1 or both more than has heretofore been done. For this reason we give here modified forms A(ii)$'$, B$'$, C$'$ of A(II); B, and C with the understanding that *when any of the conditions from* A(o) *to* C$'$ *are assumed this fact will be explicitly stated in the theorem.*

A(ii)$'$ *The product of two open sets is open; where, as defined above, an open set is either the void set or a sum of neighborhoods.*

B$'$ *For every set* E *we have* $\bar{\bar{E}} = \bar{E}$.

C$'$ *The sum of two non-dense sets is non-dense.*

It is readily seen from the conditions A that the ordinary closure $\bar{E}$ satisfies the conditions B so that when one assumes A(o), A(i), A(ii) and B$'$ it is never necessary to assume B(i) or B(ii). Also C$'$ is implied by the conditions A(o), A(i), A(ii)$'$ and B$'$. The condition A(ii)$'$ may be stated in the equivalent form:

A(ii)$'$ *If* $NN'E \neq 0$ *there is an* N'' *such that* $N'' \subset NN'$ *and* $N''E \neq 0$.

It is in this form that we shall most frequently use it. The final condition which will sometimes be used is

R *For every* N *there is an* N' *such that* $\bar{N}' \subset N$.

A set function $f(e)$ on $\mathfrak{P}(X)$ to $\mathfrak{P}(X_1)$ is said to be *almost* 1-1 in case of $f(N)f(N')$ is non-dense, whenever N, N' are neighborhoods whose product is void, and it is said to be *continuous* in case for every e in $\mathfrak{P}(X)$ and $N_1 \epsilon \mathfrak{N}_1$ satisfying the conditions $0 \neq e, f(e) \subset N_1$ there is an $N \epsilon \mathfrak{N}$ such that $Ne \neq 0$ and $f(N) \subset N_1$. The set function $f(e)$ is said to be *monotone* if $e \subset e'$ implies $f(e) \subset f(e')$. Many of the properties of set functions which can be defined by point functions are not needed and will not be assumed. Thus for example, we do not assume that $f(\sum e_n) = \sum f(e_n)$. It is assumed however in the last part of Theorem 5.1 that subsets of the image $f(E)$ of E are images of subsets of E, a property satisfied by every set function which is defined by a point function in the usual fashion.

THEOREM 5.1: *Let the space* X *with its topology* $\mathfrak{N}$ *satisfy* A *and let the space* X_1 *with its family* $\mathfrak{N}_1$ *of neighborhoods and closure operation satisfy the axioms* A, B, *and* C$'$. *Let* $f(e)$ *be a continuous monotone almost* 1-1 *set function with domain* $\mathfrak{P}(X)$ *and range*[6] $\mathfrak{P}(X_1)$ *which satisfies the condition*

(i) $$f(\mathfrak{N}) \subset \sum_{\phi} \mathfrak{B}_{\phi} II_{\phi}.$$

[6] The range is the set of functional values.

Then

(ii) *For every $N \in \mathfrak{N}$ there is an $N_1 \in \mathfrak{N}_1$ such that $f(\tilde{N}) \supset N_1$.*

If in addition it is assumed that the space X satisfies the regularity condition R, *the space X_1 satisfies the condition* B′, *and the function f satisfies the condition $f(\mathfrak{P}(E)) = \mathfrak{P}(f(E))$; then the three statements* (i), (ii), *and* (iii) *are all equivalent.*

(iii) *For every non-void open set G in X, $f(G) = G_1 + E_1$, where G_1 is a non-void open set in X_1 and E_1 is a non-dense set which is closed in $f(G)$.*

Let $\phi(N)$ on $\mathfrak{N}$ be such that $f(N) \in \mathfrak{B}_{\phi(N)}II_{\phi(N)}$ for every $N \in \mathfrak{N}$. For each N there is by Theorem 3.3 an $N_1 \in \mathfrak{N}_1$ such that

(a) $$f(N) \text{ is locally } II_{\phi(N)} \text{ on } N_1$$

(b) $$E_1 = N_1 - f(N) \in I_{\phi(N)}.$$

Let $E \subset X$ be such that $f(E) = E_1$. It will now be shown that $E \subset \tilde{N}$. If there is a point $x \in E - \tilde{N}$ then there will be a neighborhood $N(x)$ of x such that $N(x)N = 0$, $f(N(x))f(N)$ is non-dense. Since $f(x) \in f(E) = E_1 \subset N_1$ it may be supposed, using continuity and A(ii), that N^0 has been chosen so that

(c) $$f(N^0) \subset N_1, \qquad f(N^0)f(N) \text{ is non-dense}, \qquad N^0 \subset N(x).$$

Now (c) and (b) show that $f(N^0) \subset E_1 + f(N^0)f(N)$, and thus that $f(N^0) \in I_{\phi(N)}$. Hence

(d) $$\phi(N^0) < \phi(N).$$

Using Theorem 3.3 and A(ii) on the space X_1 we can find an $N_1' \in \mathfrak{N}_1$ such that $N_1' \subset N_1$ and

(a′) $$f(N^0) \text{ is locally } II_{\phi(N^0)} \text{ on } N_1',$$

(b′) $$N_1' - f(N^0) \in I_{\phi(N^0)}.$$

Since $N_1' \subset N_1$ we have by (a) and (c) the set $f(N)[N_1' - f(N^0)] = N_1'f(N) - f(N^0)f(N) \in II_{\phi(N)}$, and so $N_1' - f(N^0)$ must be a $II_{\phi(N)}$ set. This shows in view of (b′) that $\phi(N) < \phi(N^0)$, which is a contradiction to (d). Thus we can say that $E \subset \tilde{N}$. Now $N_1 \subset E_1 + f(N)$ and $E_1 = f(E) \subset f(\tilde{N})$ and so

$$N_1 \subset f(\tilde{N}) + f(N) = f(\tilde{N}).$$

This establishes the first conclusion of the theorem.

Since an open set G_1 and a non-dense set E_1 both belong to $\mathfrak{B}_\phi$ for any ϕ, finite or transfinite,[7] and for sufficiently small ϕ (ϕ finite for example) $G_1 + E_1 \in II_\phi$ we see that (iii) implies (i). Since it has already been shown that (i) implies (ii) it will be sufficient in proving the second assertion of the theorem to prove that (ii) implies (iii).

[7] Theorem 3.4 (ii). Here C(i) must be satisfied since we have assumed that $\bar{E} = \bar{\bar{E}}$ in X_1. Also we have assumed C(ii)′ in X_1.

It is clear that if X satisfies R we have for every non-void open G an $N_1 \epsilon \mathfrak{N}_1$ such that $N_1 \subset f(G)$. Let G_1 be the sum of all $N_1 \subset f(G)$ and consider the set

$$E_1 = f(G) - G_1.$$

If E_1 is not closed in $f(G)$, i.e., if $0 \neq \bar{E}_1 f(G) - E_1 \subset G_1$, there is a neighborhood $N_1 \subset G_1$ (and hence $N_1 E_1 = 0$) such that $N_1[\bar{E}_1 f(G) - E_1] \neq 0$, so that $N_1 \bar{E}_1 \neq 0$ and $N_1 E_1 = 0$, which is a contradiction to the definition of closure.

If $\bar{E}_1$ contains a neighborhood N_1 we must have $G_1 N_1 = 0$ for otherwise there would be a neighborhood N_1' such that

$$N_1' \subset N_1 G_1 \subset f(G),$$

$$N_1' \subset N_1 G_1 \subset \bar{E}_1,$$

$$N_1' \subset \bar{E}_1 f(G) = E_1,$$

which is a contradiction to the definition of E_1. Hence we can say that if $\bar{E}_1 \supset N_1$ then $N_1 G_1 = 0$. Let $N_1 \subset \bar{E}_1$ so that $N_1 E_1 \neq 0$. Pick $e \subset G$ so that $f(e) = N_1 E_1$. By continuity there is an $N' \epsilon \mathfrak{N}$ such that $N'e \neq 0, f(N') \subset N_1$, and since $e \subset G$ we have $0 \neq N'e = N'Ge$. There is then an $N'' \subset N'G$ and so $f(N'') \subset f(N') \subset N_1$. But there is an N_1'' such that

$$N_1'' \subset f(N'') \subset f(G), \qquad N_1'' \subset f(N'') \subset N_1,$$

and so $N_1'' \subset G_1 N_1 = 0$, a contradiction to A(i). Thus $\bar{E}_1$ contains no neighborhood and the proof of Theorem 5.1 is complete.

There are functions other than almost 1-1 functions for which conclusion (iii) of Theorem 5.1 holds, at least when applied to metric spaces. To illustrate a class of such functions we state the following definition. A function f defined on a metric space X and with values in an arbitrary class is said to be *distributive* in case for every positive number p there is a positive number q such that if x, x', N have the properties[8]

(a) $$x \epsilon N \epsilon \mathfrak{N}, \text{ diam } N < q, \qquad f(x) = f(x')$$

there will be an N' such that

(b) $$x' \epsilon N' \epsilon \mathfrak{N}, \text{ diam } N' < p, \qquad f(N) = f(N').$$

Thus any 1-1 function will be distributive. Also any linear mapping between linear spaces will be distributive providing the metric in the linear spaces satisfies the condition $(x, y) = (x - y, 0)$. Also any homomorphic mapping between metric groups will be distributive in case the groups have a left (or right) invariant metric.

THEOREM 5.2: *Let X be metric and X_1 satisfy*[9] A, B′, *and* C′. *Let f be a*

[8] The family $\mathfrak{N}$ in the case of a metric space is taken as the family of all open spheres.

[9] Or since C(ii)′ is implied by A, A(ii)′, B′ we may assume instead that X_1 satisfies A, A(ii)′, B′.

continuous distributive function with domain X and range X_1 which satisfies the condition

$$f(\mathfrak{N}) \subset \sum_{\phi} \mathfrak{B}_{\phi} II_{\phi}.$$

Then for every non-void open set G in X we have $f(G) = G_1 + E_1$ where G_1 is a non-void open set in X_1 and E_1 is a non-dense set which is closed in $f(G)$.

For each x_1 in X_1 let E_{x_1} be the set of all $x \in X$ for which $x_1 = f(x)$. The sets E_{x_1} $(x_1 \in X_1)$ are then disjoint subsets of X and there is by the axiom of choice a set $X_0 \subset X$ such that for each x_1 the set $E_{x_1}X_0$ consists of a single point. The function $f_0(x_0) = f(x_0)$ defined on X_0 is then a 1-1 map of X_0 into X_1. Any set of the form $f_0^{-1}f(N)$ where $N \in \mathfrak{N}$ (i.e., N is an open sphere) will be called a neighborhood N_0 in the space X_0. For each N in $\mathfrak{N}$ there will be then one and only one neighborhood $N_0 \subset X_0$ such that $f_0(N_0) = f(N) = f(N_0)$ but for a given N_0 there may be many neighborhoods N for which this equality holds. If $\mathfrak{N}_0$ is the family of all neighborhoods $N_0 \subset X_0$ then it is clear that $f_0(\mathfrak{N}_0) \subset \sum_{\phi} \mathfrak{B}_{\phi} II_{\phi}$, and $f_0(x_0)$ is a continuous mapping of X_0 into X_1. We wish now to establish certain properties of the $\mathfrak{N}_0$ topology. It is clear that A(o), A(i) are satisfied. The topology $\mathfrak{N}_0$ also satisfies the condition A(ii)′ for if $x_0 \in N_0N_0'$ where $f(N_0) = f(N), f(N_0') = f(N')$, there will be points x, x' in N, N' respectively such that $f(x_0) = f(x) = f(x')$. In view of the distributive character of f there will be neighborhoods M, M' of x, x' respectively ($M, M' \in \mathfrak{N}$) such that $M \subset N$, $M' \subset N'$ and $f(M) = f(M')$. Thus the neighborhood $N_0'' = f_0^{-1}f(M) = f_0^{-1}f(M')$ has the properties $N_0'' \in \mathfrak{N}_0$, $x_0 \in N_0'' \subset N_0N_0'$. We wish next to show that the topology $\mathfrak{N}_0$ satisfies R. Let $N_0 = f_0^{-1}f(N)$. Let x be a point of N and $p > 0$ be such that the open sphere $S(x, 2p)$ with center x and radius $2p$ is entirely within N. Let $M = S(x, p)$ so that $\bar{M} \subset N$. Define $M_0 = f_0^{-1}f(M)$, so that $M_0 \in \mathfrak{N}_0$. Let $x_0 \in \bar{M}_0$ where the closure is taken with respect to the $\mathfrak{N}_0$ topology; i.e., every N_0 of x_0 intersects M_0. Let $N(x_0) \in \mathfrak{N}$ with diameter $N(x_0) < q$. Let $N_0(x_0) = f_0^{-1}f(N(x_0))$ so that $N_0(x_0) \in \mathfrak{N}_0$ and $x_0 \in N_0(x_0)$. Then $N_0(x_0)M_0 \neq 0$, and hence $f(N(x_0))f(M) = f_0(N_0(x_0))f_0(M_0) \neq 0$. There is then a $y \in N(x_0)$ and a $z \in M$ such that $f(y) = f(z)$. Let $N' \in \mathfrak{N}$ be a neighborhood of z with diameter $N' < p$ and such that $f(N') = f(N(x_0))$. Now for any $x' \in N'$ we have $(x', x) \leqq (x', z) + (z, x) < p + p = 2p$ so that $N' \subset N$. This shows that $x_0 \in f_0^{-1}f(N(x_0)) = f_0^{-1}f(N') \subset f_0^{-1}f(N) = N_0$. Since x_0 was an arbitrary point in the $\mathfrak{N}_0$ closure of M_0 we have $\bar{M}_0 \subset N_0$ and the topology $\mathfrak{N}_0$ satisfies the condition R. The desired conclusion now follows immediately by applying Theorem 5.1 to the set function determined by f_0 on X_0 to X_1.

Our attention will now be turned to the problem of finding when the hypothesis

$$f(\mathfrak{N}) \subset \sum_{\phi} \mathfrak{B}_{\phi} II_{\phi}$$

holds for every continuous function f which takes neighborhoods into sets which are not non-dense. We have not been able to settle this question for any case

except the one where X is a complete metric space. One condition sufficient for the validity of $f(\mathfrak{N}) \subset \sum_\phi \mathfrak{B}_\phi II_\phi$ is of the form $s(x) < c(x)$ where $s(x)$ and $c(x)$ are numerical invariants of the point x called *the separability character of the space X at the point x* and *the category character of the space X at the point x*, respectively, which we now proceed to define. If a set E contains no points which are isolated in X then there is a smallest cardinal number $c(E)$, called *the category character of E*, such that E is the sum of $c(E)$ non-dense sets. If E contains a point x which is isolated in X (i.e., the set consisting of the single point x is an open set) then we write $c(x) = u$, and read this as $c(E)$ is undefined. If for every neighborhood N of x we have $c(N) = u$ then we put $c(x) = u$ and otherwise define $c(x)$ as the greatest lower bound of the numbers $c(N)$ where N varies over the neighborhoods of x. The number $c'(x)$ is then the first ordinal in the set $c'(N)$ and the value $c(x)$ is attained by the function $c(N)$. By the separability character of an open set G in X we shall mean the smallest cardinal $s(G)$ such that there exist open sets

$$\text{(a)} \qquad G_1, G_2, \cdots, G_\alpha, \cdots \qquad \alpha < s'(G)$$

forming a basis for G; i.e., every open subset of G is the sum of the G_α's contained in it, and $G_\alpha \subset G$. The number $s(x)$ is then defined as the greatest lower bound of the numbers $s(N)$ where N varies over the neighborhoods of x, and as before it is seen that $s'(x)$ is the first ordinal in the set $s'(N)$ and that $s(x)$ is attained by $s(N)$ for some neighborhood N of x. It is clear then that *for every x in X there is a neighborhood N_0 of x such that if N is any neighborhood of x which is contained in N_0 we have $s(x) = s(N)$, $c(x) = c(N)$*. We shall agree to write $s(x) < c(x)$ for every x with $c(x) = u$.

THEOREM 5.3: *If X is a metric space which is locally complete and locally separable then $s(x) < c(x)$ for every x in X.*

This is nothing more than the well known category theorem of Baire applied to the case in hand.

Unfortunately the locally separable spaces are, as far as we know, the only ones which satisfy the condition $s(x) < c(x)$, and so what we have to say is essentially applicable only to these spaces. The range space X_1, however, will be left quite arbitrary. A space is said to satisfy the condition (H) in case

(H) *for every pair x, x' of points with $x \neq x'$ there are neighborhoods N, N' such that $x \in N$, $x' \in N'$, $NN' = 0$.*

THEOREM 5.4: *Let X be an arbitrary*[10] *complete metric space and G an open set of X. Let $f(x)$ be a continuous function on X to X_1 where the space X_1 satisfies the conditions* A, B, C′ *and* H. *Then*

$$f(G) \in \mathfrak{B}_{s(G)}.$$

If $s(G)$ is finite G consists of a finite number of isolated points and since the theorem is easily verified in this case we shall assume that $s(G) \geqq \aleph_0$.

[10] I.e., we do not assume local separability or $s(x) < c(x)$.

continuous distributive function with domain X and range X_1 which satisfies the condition

$$f(\mathfrak{N}) \subset \sum_{\phi} \mathfrak{B}_{\phi} II_{\phi}.$$

Then for every non-void open set G in X we have $f(G) = G_1 + E_1$ where G_1 is a non-void open set in X_1 and E_1 is a non-dense set which is closed in $f(G)$.

For each x_1 in X_1 let E_{x_1} be the set of all $x \epsilon X$ for which $x_1 = f(x)$. The sets E_{x_1} $(x_1 \epsilon X_1)$ are then disjoint subsets of X and there is by the axiom of choice a set $X_0 \subset X$ such that for each x_1 the set $E_{x_1}X_0$ consists of a single point. The function $f_0(x_0) = f(x_0)$ defined on X_0 is then a 1-1 map of X_0 into X_1. Any set of the form $f_0^{-1}f(N)$ where $N \epsilon \mathfrak{N}$ (i.e., N is an open sphere) will be called a neighborhood N_0 in the space X_0. For each N in $\mathfrak{N}$ there will be then one and only one neighborhood $N_0 \subset X_0$ such that $f_0(N_0) = f(N) = f(N_0)$ but for a given N_0 there may be many neighborhoods N for which this equality holds. If $\mathfrak{N}_0$ is the family of all neighborhoods $N_0 \subset X_0$ then it is clear that $f_0(\mathfrak{N}_0) \subset \sum_{\phi} \mathfrak{B}_{\phi} II_{\phi}$, and $f_0(x_0)$ is a continuous mapping of X_0 into X_1. We wish now to establish certain properties of the $\mathfrak{N}_0$ topology. It is clear that A(o), A(i) are satisfied. The topology $\mathfrak{N}_0$ also satisfies the condition A(ii)$'$ for if $x_0 \epsilon N_0N_0'$ where $f(N_0) = f(N)$, $f(N_0') = f(N')$, there will be points x, x' in N, N' respectively such that $f(x_0) = f(x) = f(x')$. In view of the distributive character of f there will be neighborhoods M, M' of x, x' respectively (M, $M' \epsilon \mathfrak{N}$) such that $M \subset N$, $M' \subset N'$ and $f(M) = f(M')$. Thus the neighborhood $N_0'' = f_0^{-1}f(M) = f_0^{-1}f(M')$ has the properties $N_0'' \epsilon \mathfrak{N}_0$, $x_0 \epsilon N_0'' \subset N_0N_0'$. We wish next to show that the topology $\mathfrak{N}_0$ satisfies R. Let $N_0 = f_0^{-1}f(N)$. Let x be a point of N and $p > 0$ be such that the open sphere $S(x, 2p)$ with center x and radius $2p$ is entirely within N. Let $M = S(x, p)$ so that $\bar{M} \subset N$. Define $M_0 = f_0^{-1}f(M)$, so that $M_0 \epsilon \mathfrak{N}_0$. Let $x_0 \epsilon \tilde{M}_0$ where the closure is taken with respect to the $\mathfrak{N}_0$ topology; i.e., every N_0 of x_0 intersects M_0. Let $N(x_0) \epsilon \mathfrak{N}$ with diameter $N(x_0) < q$. Let $N_0(x_0) = f_0^{-1}f(N(x_0))$ so that $N_0(x_0) \epsilon \mathfrak{N}_0$ and $x_0 \epsilon N_0(x_0)$. Then $N_0(x_0)M_0 \neq 0$, and hence $f(N(x_0))f(M) = f_0(N_0(x_0))f_0(M_0) \neq 0$. There is then a $y \epsilon N(x_0)$ and a $z \epsilon M$ such that $f(y) = f(z)$. Let $N' \epsilon \mathfrak{N}$ be a neighborhood of z with diameter $N' < p$ and such that $f(N') = f(N(x_0))$. Now for any $x' \epsilon N'$ we have $(x', x) \leqq (x', z) + (z, x) < p + p = 2p$ so that $N' \subset N$. This shows that $x_0 \epsilon f_0^{-1}f(N(x_0)) = f_0^{-1}f(N') \subset f_0^{-1}f(N) = N_0$. Since x_0 was an arbitrary point in the $\mathfrak{N}_0$ closure of M_0 we have $\tilde{M}_0 \subset N_0$ and the topology $\mathfrak{N}_0$ satisfies the condition R. The desired conclusion now follows immediately by applying Theorem 5.1 to the set function determined by f_0 on X_0 to X_1.

Our attention will now be turned to the problem of finding when the hypothesis

$$f(\mathfrak{N}) \subset \sum_{\phi} \mathfrak{B}_{\phi} II_{\phi}$$

holds for every continuous function f which takes neighborhoods into sets which are not non-dense. We have not been able to settle this question for any case

except the one where X is a complete metric space. One condition sufficient for the validity of $f(\mathfrak{N}) \subset \sum_\phi \mathfrak{B}_\phi II_\phi$ is of the form $s(x) < c(x)$ where $s(x)$ and $c(x)$ are numerical invariants of the point x called *the separability character of the space X at the point x* and *the category character of the space X at the point x*, respectively, which we now proceed to define. If a set E contains no points which are isolated in X then there is a smallest cardinal number $c(E)$, called *the category character of E*, such that E is the sum of $c(E)$ non-dense sets. If E contains a point x which is isolated in X (i.e., the set consisting of the single point x is an open set) then we write $c(x) = u$, and read this as $c(E)$ is undefined. If for every neighborhood N of x we have $c(N) = u$ then we put $c(x) = u$ and otherwise define $c(x)$ as the greatest lower bound of the numbers $c(N)$ where N varies over the neighborhoods of x. The number $c'(x)$ is then the first ordinal in the set $c'(N)$ and the value $c(x)$ is attained by the function $c(N)$. By the separability character of an open set G in X we shall mean the smallest cardinal $s(G)$ such that there exist open sets

$$G_1, G_2, \cdots, G_\alpha, \cdots \qquad \alpha < s'(G) \tag{a}$$

forming a basis for G; i.e., every open subset of G is the sum of the G_α's contained in it, and $G_\alpha \subset G$. The number $s(x)$ is then defined as the greatest lower bound of the numbers $s(N)$ where N varies over the neighborhoods of x, and as before it is seen that $s'(x)$ is the first ordinal in the set $s'(N)$ and that $s(x)$ is attained by $s(N)$ for some neighborhood N of x. It is clear then that *for every x in X there is a neighborhood N_0 of x such that if N is any neighborhood of x which is contained in N_0 we have $s(x) = s(N)$, $c(x) = c(N)$.* We shall agree to write $s(x) < c(x)$ for every x with $c(x) = u$.

THEOREM 5.3: *If X is a metric space which is locally complete and locally separable then $s(x) < c(x)$ for every x in X.*

This is nothing more than the well known category theorem of Baire applied to the case in hand.

Unfortunately the locally separable spaces are, as far as we know, the only ones which satisfy the condition $s(x) < c(x)$, and so what we have to say is essentially applicable only to these spaces. The range space X_1, however, will be left quite arbitrary. A space is said to satisfy the condition (H) in case

(H) *for every pair x, x' of points with $x \neq x'$ there are neighborhoods N, N' such that $x \in N$, $x' \in N'$, $NN' = 0$.*

THEOREM 5.4: *Let X be an arbitrary*[10] *complete metric space and G an open set of X. Let $f(x)$ be a continuous function on X to X_1 where the space X_1 satisfies the conditions* A, B, C′ *and* H. *Then*

$$f(G) \in \mathfrak{B}_{s(G)}.$$

If $s(G)$ is finite G consists of a finite number of isolated points and since the theorem is easily verified in this case we shall assume that $s(G) \geq \aleph_0$.

[10] I.e., we do not assume local separability or $s(x) < c(x)$.

Let G be an open set and suppose that the sets G_α in the series (a) form a basis for G. Let H be an open set such that $\bar{H} \subset G$ and define $H_\alpha = HG_\alpha$ so that the sequence

$$\text{(b)} \qquad H_1, H_2, \cdots, H_\alpha, \cdots \qquad \alpha < s'(G)$$

forms a basis for H. For certain points x in X there will be at least one denumerable sequence ζ_n of ordinals $< s'(G)$ for which

$$\text{(c)} \qquad x \epsilon \bar{H}_{\zeta_n}, \qquad \bar{H}_{\zeta_n} \subset S_c(x, 1/n), \qquad n = 1, 2, \cdots,$$

where $S_c(x, 1/n)$ is the closed sphere with center x and radius $1/n$ (i.e., the set of all y in X such that $(x, y) \leqq 1/n$; this is not necessarily the closure of the set $S(x, 1/n)$ of y for which $(x, y) < 1/n$). For example if x is a point of H, there is for each $n = 1, 2, \cdots$ a sphere $S(x, 2r_n) \subset HS_c(x, 1/n)$ and since (b) is a basis for H there is an ordinal $\zeta_n < s'(G)$ such that $x \epsilon \bar{H}_{\zeta_n} \subset S(x, r_n)$ and so

$$x \epsilon \bar{H}_{\zeta_n} \subset \overline{S(x, r_n)} \subset S(x, 2r_n) \subset S_c(x, 1/n).$$

There may be points besides those in H for which such sequences exist. Let H^* be the set of all points x in X for which at least one sequence satisfying (c) exists. Then, obviously,

$$\text{(d)} \qquad H \subset H^* \subset \bar{H} \subset G.$$

Let Z be the space, discussed in §4, of all denumerable sequences $\{\zeta_n\}$ of ordinals $\zeta_n < s'(G)$. For every x in H^* we define $Z(x)$ as the family of all $\{\zeta_n\} \epsilon Z$ for which (c) holds. Since for $\mathfrak{z} = \{\zeta_n\} \epsilon Z(x)$ we have

$$\text{diameter} \prod_{n=1}^{\infty} \bar{H}_{\zeta_n} \leqq \text{diameter } \bar{H}_{\zeta_m} \leqq \text{diameter } S_c(x, 1/n) \leqq 2/m,$$

the set $\prod_{n=1}^{\infty} \bar{H}_{\zeta_n}$ reduces to the single point x. This shows that for every x and y in H^*, with $x \neq y$ we have $Z(x)Z(y) = 0$. It will now be shown that the set

$$\text{(e)} \qquad Z(H^*) = \sum_{x \epsilon H^*} Z(x)$$

is a closed set in Z. Suppose $\mathfrak{z}_n = \{\zeta_p^n\}$ is a sequence of points in $Z(H^*)$, with $\mathfrak{z}_n \rightarrow \mathfrak{z}_0 = \{\zeta_p^0\}$. From the definition of the metric in Z one can find a sequence $m_n \rightarrow \infty$ such that

$$\text{(f)} \qquad \zeta_p^0 = \zeta_p^n \qquad p = 1, 2, \cdots, m_n.$$

Since $x, y \epsilon H^*$, $x \neq y$ implies $Z(x)Z(y) = 0$ we have x_n in H^* uniquely determined by the condition $\mathfrak{z}_n \epsilon Z(x_n)$, and thus

$$\text{(g)} \qquad x_n \epsilon \bar{H}_{\zeta_p^n} \subset S_c(x_n, 1/p) \qquad p = 1, 2, \cdots.$$

Thus by (f) we have

$$\text{(h)} \qquad x_n \epsilon \bar{H}_{\zeta_p^0} \subset S_c(x_n, 1/p) \qquad p = 1, 2, \cdots, m_n.$$

It is obviously no loss of generality to assume that the sequence m_n is monotone and thus from (h) we have for $r \leqq n$ the set $\prod_{p=1}^{m_r} \bar{H}_{\zeta_p^0}$ containing both x_r and x_n, and so for $r \leqq n$, it follows also from (h) that

$$(x_r, x_n) \leqq \text{diameter} \prod_{p=1}^{m_r} \bar{H}_{\zeta_p^0} \leqq \text{diameter } \bar{H}_{\zeta_{m_r}^0} \leqq \text{diameter } S_c(x_n, 1/m_r) \leqq 2/m_r.$$

This shows that the sequence $\{x_n\}$ is a Cauchy sequence. Let $x_0 = \lim x_n$. All that remains to be shown is that $x_0 \in H^*$ and that $z_0 \in Z(x_0)$. These two statements can be verified simultaneously by showing that

$$\text{(i)} \qquad x_0 \in \bar{H}_{\zeta_p^0} \subset S_c(x_0, 1/p) \qquad p = 1, 2, \cdots .$$

For a fixed p we have $m_n > p$ for sufficiently large n, and thus (by (h)) $x_n \in \bar{H}_{\zeta_p^0}$ if n is large. This shows that $x_0 \in \bar{H}_{\zeta_p^0}$ for every integer p. Also by (h) for every $y \in \bar{H}_{\zeta_p^0}$, $(y, x_n) \leqq 1/p$ if $m_n \geqq p$ and thus $(y, x_0) \leqq 1/p$. This establishes (i) and completes the proof that $Z(H^*)$ is a closed set.

For every integer $n = 1, 2, \cdots$ and every $\mathfrak{z} = \{\zeta_p\}$ in the closed set $Z(H^*)$ put

$$K_{\zeta_1 \cdots \zeta_n} = \prod_{p=1}^{n} \widehat{f(H_{\zeta_p})}$$

so that

$$\prod_{n=1}^{\infty} K_{\zeta_1 \cdots \zeta_n} = \prod_{n=1}^{\infty} \widehat{f(H_{\zeta_n})}.$$

Suppose $\mathfrak{z}$ corresponds to $x \in H^*$, i.e., $\mathfrak{z} \in Z(x)$, then using (c) and the continuity of f we have

$$f(x) \in f(\bar{H}_{\zeta_n}) \subset \widehat{f(H_{\zeta_n})} \qquad n = 1, 2, \cdots ,$$

and thus $f(x) \in \prod_{n=1}^{\infty} \widehat{f(H_{\zeta_n})}$. Let $x_1 = f(x)$ and suppose that there is another point $x_1' \neq x_1$ also in $\prod_{n=1}^{\infty} \widehat{f(H_{\zeta_n})}$. It is readily seen from the condition H which was assumed to hold in X_1 that there is a neighborhood N_1' containing x_1' and such that $x_1 - \in \bar{N}_1'$. Since $x_1' \in \widehat{f(H_{\zeta_n})}$ for every n we have $N_1' f(H_{\zeta_n}) \neq 0$ and hence points x_n for which

$$f(x_n) \in N_1', \qquad x_n \in H_{\zeta_n}, \qquad n = 1, 2, \cdots .$$

Since

$$x_n \in H_{\zeta_n} \subset \bar{H}_{\zeta_n} \subset S_c(x, 1/n) \qquad n = 1, 2, \cdots ,$$

the point $x_n \to x$. Since $x_1 - \in \bar{N}_1'$ there is a neighborhood N_1 of x_1 such that $N_1 N_1' = 0$ and by continuity we must have $f(x_n) \in N_1$ for all sufficiently large n which is a contradiction to the fact that $f(x_n) \in N_1'$. Thus $\prod_{n=1}^{\infty} \widehat{f(H_{\zeta_n})}$ reduces to the single point $f(x)$ and thus

$$f(H^*) = \sum \prod_{n=1}^{\infty} K_{\zeta_1 \cdots \zeta_n} = A_{Z(H^*)}[K], \qquad \text{summed for } \mathfrak{z} \in Z(H^*).$$

It is readily shown that $\overline{f(H_{\zeta_n})}$ is the complement of an open set and thus by Theorem 3.4 the kernel $K_{\zeta_1 \cdots \zeta_n}$ which is defined over $Z(H^*)$ has its values in $\mathfrak{B}_{s(G)}$. By Theorem 4.1 the set $f(H^*)$ is a $\mathfrak{B}_{s(G)}$ set. To each x in G assign an integer $n(x)$ such that $S(x, 2/n(x)) \subset G$, and let

$$H_m = \sum S(x, 1/n(x)),$$

where the sum is taken over all x in G for which $n(x) = m$. Let $x_p \in H_m$ and $x_p \to x_0$. Then for each p there is an x^p with

$$n(x^p) = m, \qquad x_p \in S(x^p, 1/m).$$

For sufficiently large p, $(x_p, x_0) < 1/2m$ and thus for large p

$$(x^p, x_0) \leqq (x^p, x_p) + (x_p, x_0) < 1/m + 1/2m < 2/m,$$

so that

$$x_0 \in S(x^p, 2/m) = S(x^p, 2/n(x^p)) \subset G.$$

Thus $\bar{H}_m \subset G$, and using (d),

$$H_m \subset H_m^* \subset \bar{H}_m \subset G.$$

From the definition of H_m we see that $\sum_{m=1}^{\infty} H_m = \sum_{m=1}^{\infty} H_m^* = G$ and thus the set

$$f(G) = \sum_{m=1}^{\infty} f(H_m^*)$$

is, by Theorem 3.4, a $\mathfrak{B}_{s(G)}$ set. This completes the proof of Theorem 5.4.

Now let $x_1 = f(x)$ be a continuous mapping of one topological space X onto another space X_1 and suppose that in both spaces $\bar{E} = \bar{\bar{E}}$. Suppose f is such that for every neighborhood N the set $f(N)$ is not a non-dense set. Then since $f(\bar{E}) \subset \overline{f(E)}$ we see that E must be non-dense if $f(E)$ is. Likewise if $f(E) \in I_\phi$, i.e.,

$$f(E) = \sum_{\alpha < \phi'} E_1^\alpha, \qquad E_1^\alpha \text{ non-dense},$$

we may take $E^\alpha \subset E$ so that $f(E^\alpha) = E_1^\alpha$ and $E = \sum_{\alpha<\phi'} E^\alpha$. Since $f(E^\alpha) = E_1^\alpha$ is non-dense so is E^α and thus $E \in I_\phi$. In other words, for continuous functions which take neighborhoods into sets which are not non-dense we must have

$$f(N) \in II_\phi \qquad\qquad \phi < c(N).$$

As pointed out above there is for every x a neighborhood N_0 containing x such that for every neighborhood N of x with $N \subset N_0$ we have $s(x) = s(N)$, $c(x) = c(N)$. For a fixed x let $\mathfrak{N}_x$ be the set of all N in $\mathfrak{N}$ for which $x \in N \subset N_0$, and let $\mathfrak{N}' = \sum_x \mathfrak{N}_x$. The family $\mathfrak{N}'$ then determines (assuming A(ii)′) the same topology (i.e., the same notion of open set) as that determined by the larger family $\mathfrak{N}$. We can say then that every $N \in \mathfrak{N}'$ contains at least one point x for which $s(x) = s(N)$, $c(x) = c(N)$. These equalities of course do not

have to hold for all points in N and thus it is not true that for every neighborhood $N \epsilon \mathfrak{N}'$ of a point x we have $s(x) = s(N)$, $c(x) = c(N)$. Suppose the locally complete metric space X has the property $s(x) < c(x)$ and that the family $\mathfrak{N}$ (i.e., the set of open spheres) has been replaced by the set of $N' \epsilon \mathfrak{N}'$ for which $\bar{N}'$ is complete. Let $f(x)$ be a continuous function on X having the property that open sets transform into sets which are not non-dense. Then for every $N' \epsilon \mathfrak{N}'$ there is at least one point $x \epsilon N'$ such that $s(N') = s(x) < c(x) = c(N')$ and so by the preceding remarks we have $f(N') \epsilon II_{s(N')}$. But by Theorem 5.4, when proper restrictions are placed on X_1, we have $f(N') \epsilon \mathfrak{B}_{s(N')}$ so that

$$f(\mathfrak{N}') \subset \sum_{\phi} \mathfrak{B}_{\phi} II_{\phi}.$$

The following theorem then follows immediately from Theorems 5.2, 5.3, 5.4.

THEOREM 5.5: *Let X be a locally complete and locally separable metric space, and let X_1 be a Hausdorff space.*[11] *Let $f(x)$ be a continuous distributive function which maps X into the entire space X_1 in such a way that non-void open sets in X are transformed into sets which are not non-dense.*[12] *Then for every non-void open set G in X we have $f(G) = G_1 + E_1$, where G_1 is a non-void open set in X_1 and E_1 is a non-dense set which is closed in $f(G)$.*

6. Locally homogeneous spaces

In this section we shall make an application of the results of §5 to a situation somewhat more general than an additive function on a group.

6.1. *Preliminary lemmas.* All of the spaces considered in §6 are neighborhood spaces in the sense of Hausdorff, i.e., we have a family $\mathfrak{N}$ of non-void sets N, each of which is called a neighborhood of any one of its points, and the following assumptions are made.

(i) If N_x, N'_x are neighborhoods of x then there is a neighborhood M_x of x with $M_x \subset N_x N'_x$.

(ii) If $x \neq y$ there are neighborhoods N_x, N_y of x and y respectively with $N_x N_y = 0$.

Any sum of neighborhoods is called an open set and the family of open sets will be denoted by $\mathfrak{G}$. The closure $\bar{E}$ of a set E is the set of all x for which $N_x E \neq 0$ for every neighborhood N_x of x. When we are dealing with two spaces X and X_1 the neighborhoods and open sets in X_1 will be distinguished from those in X by the subscript 1. Besides being a Hausdorff space we assume that the space X is *locally homogeneous* in the following sense.

(iii) $\mathfrak{H}$ is a family of homeomorphisms $y = h(x)$. Each of these homeomorph-

[11] I.e., a space satisfying A(o), A(i), A(ii)′, B′ and H.

[12] This last assumption and the assumption of local separability (or $s(x) < c(x)$) may of course both be replaced by the single assumption that for every neighborhood N of a point x we have $f(N) \epsilon II_{s(x)}$.

isms $h \in \mathfrak{H}$ has for its domain of definition $D(h)$, an open subset of X. The range $R(h)$, i.e., the set of functional values of h is also an open subset of X.
(iv) The identity function $x = h(x)$ defined throughout X is an element of $\mathfrak{H}$.
(v) For every pair a, b of points in X there is one and only one $h_{ab} \in \mathfrak{H}$ having the properties $a \in D(h_{ab})$, $b \in R(h_{ab})$, $b = h_{ab}(a)$. We shall sometimes write $(a, b; x)$ for $h_{ab}(x)$.
(vi) The function $(x, y; z)$ is continuous in its arguments; and by this we mean that if $z \in D(h_{xy})$ and N is a neighborhood of $(x, y; z)$ then there are neighborhoods N_x, N_y, N_z of x, y, z respectively such that if $x' \in N_x$, $y' \in N_y$, $z' \in N_z$ we have $z' \in D(h_{x'y'})$ and $(x', y'; z') \in N$, i.e., $(N_x, N_y; N_z) \subset N$.

It is clear that Hausdorff linear spaces, groups, or fields of point sets are locally homogeneous spaces as defined above. For linear spaces the family $\mathfrak{H}$ may be taken as all functions of the form $a + x$; for groups all functions of the form ax and for fields of point sets all functions of the form $a + x - ax$. In each of these examples we have $D(h) = R(h) = X$ for every $h \in \mathfrak{H}$. Perhaps the simplest example of a locally homogeneous space for which $D(h) \neq X$ is the semi-group X composed of all real positive x, where the family $\mathfrak{H}$ is taken as all functions of the form $h(x) = a + x$ where a is positive or negative. Here the domain $D(h)$ is the set of all positive $x > -a$ and the range $R(h)$ is the set of all positive $x > a$. It is desirable to consider such examples as the last one mentioned from an abstract point of view for there are many places in analysis where a homomorphic mapping from such a semi-group onto the ring of bounded linear operators on a Banach space presents itself and in such a way that it cannot be extended to the whole group.[13]

The case of the positive real axis is an instance of the following general type of locally homogeneous space. Let G be a Hausdorff group (written additively) and X an open subset of G (perhaps not containing the identity and thus not a sub-group) such that $a + b \in X$ whenever $a, b \in X$. Let the neighborhoods in X be those neighborhoods in G which are contained in X. Suppose that X is such that $(a + X)X \neq 0$ for every a in G. Let $\mathfrak{H}$ be the family of all functions of the form $h(x) = a + x$ where $a \in G$. The domain $D(h)$ is then $X(-a + X)$ and the range $R(h) = a + X(-a + X)$.

LEMMA 1: *The space X is regular.*[14]

Let x be a point of the neighborhood N. Since $x = (x, x; x)$ we have in view of (i) and (vi) a neighborhood M of x such that $(M, M; M) \subset N$. Let z be a point of $\bar{M}$. In view of (iv) and (v) $h_{zz} = h_{xx}$ and thus $x = (z, z; x) \in M$. There will then be a neighborhood N_z of z such that $(N_z, N_z; x) \subset M$. Since $z \in \bar{M}$ there

[13] The reader is referred in this connection to several papers of E. Hille. For example, *Notes on linear transformations. I*, Trans. Amer. Math. Soc., 39(1936), pp. 131–153; *On semi-groups of transformations in Hilbert space*, Proc. Nat. Acad. Sci., 24(1938), 159–161; *Notes on linear transformations. II*, Annals of Math. (2) 40(1939), pp. 1–47.

[14] I.e., for every neighborhood N of a point x there is a neighborhood M of x such that $\bar{M} \subset N$. See A. Kolmogoroff, Studia Math., vol. 5(1934), p. 31, where this property is proved for groups.

is a point $y \in MN_z$. Since z and y are both in N_z we have $(y, z; x) \in M$. Now if $u = (y, z; x)$ we have by (v) $h_{yz} = h_{xu}$ and thus $z = (y, z; y) = (x, (y, z; x); y) \in N$.

LEMMA 2: *For any pair x, x_0 there is a neighborhood N_x of x such that $N_x \subset D(h_{x'x_0})$ for every $x' \in N_x$.*

From (ii) we see that there is at least one neighborhood of every point. The desired conclusion follows immediately from (i) and (vi).

LEMMA 3: *If N and N' are neighborhoods of a point x and if x_0 is a point of N then there is a neighborhood N'' of the point x with $N'' \subset NN'$ and such that $h(N'') \subset N$ for every $h \in \mathfrak{H}$ with $N'' \subset D(h)$ and $x_0 \in h(N'')$.*

Since $x_0 = (x, x_0; x) \in N$ there is by (vi) a neighborhood N'' of x with $N'' \subset NN'$ such that $N'' \subset D(h_{x'x_0})$ for every $x' \in N''$ and such that $(N'', x_0; N'') \subset N$. Suppose $h_{ab} = h \in \mathfrak{H}$ is such that $N'' \subset D(h)$, $x_0 \in h(N'')$. Then for some $x' \in N''$ we have $x_0 = (a, b; x')$ and thus by (v) we have $h = h_{ab} = h_{x'x_0}$ and so $h(x'') = (x', x_0; x'') \in N$ for every $x'' \in N''$. Thus $h(N'') \subset N$.

A function f defined on X and with values in an arbitrary Hausdorff space X_1 is said to be *$\mathfrak{H}$-distributive* in case whenever $E \subset X$ is such that $f(E)$ is a neighborhood in X_1 then $f(h(E))$ is open for every $h \in \mathfrak{H}$ with $E \subset D(h)$. An additive (not necessarily continuous) operation between linear spaces is an example of such a distributive function.

LEMMA 4: *Let X_1 be a Hausdorff space and f continuous and $\mathfrak{H}$-distributive from X to X_1. Then if for every neighborhood N in X, $f(N)$ has an interior point, the function f will be an interior transformation, i.e., f transforms open sets into open sets.*

Let N be a neighborhood of x_0. We shall show that $f(N)$ contains a neighborhood of $x_1 = f(x_0)$. Let N_1, $E \subset N$ be such that $f(N) \supset N_1 = f(E)$. Fix x in E and, using continuity together with Lemma 2, pick a neighborhood N' of x such that $f(N') \subset N_1$ and $N' \subset D(h_{x'x_0})$ for every point x' of N'. Using Lemma 3 choose $N'' \subset NN'$ with x in N'' and such that $h(N'') \subset N$ for every $h \in \mathfrak{H}$ with $N'' \subset D(h)$ and $x_0 \in h(N'')$. Pick N_1'' so that $f(N'') \supset N_1''$, and $x'' \in N''$, $x_1'' \in N_1''$ so that $x_1'' = f(x'')$. Let $h = h_{x''x_0}$, then since $x'' \in N'' \subset N'$ we have $N'' \subset D(h)$ and $x_0 \in h(N'')$ and thus $h(N'') \subset N$. Now there is a set $E'' \subset N''$ such that $x'' \in E''$, $f(E'') = N_1''$ and thus, since f is $\mathfrak{H}$-distributive, $f(h(E''))$ is open. Hence we have $x_1 = f(x_0) = f(h(x'')) \in f(h(E'')) \subset f(h(N'')) \subset f(N)$. Since $f(h(E''))$ is open $f(N)$ contains a neighborhood of x_1.

6.2. *The equivalence of comparable topologies in X.* It is known that if a linear space is of type (F) under each of two distances and if convergence in one of the two metrics implies convergence in the other then the two metrics are equivalent and define the same topology.[15] This, of course, is equivalent to saying that a 1-1 continuous linear map of one (F) space onto the whole of another is necessarily a homeomorphism. As mentioned in the introduction the same theorem has been proved for groups but under the assumption that each of the distance

[15] S. Banach, *Théorie des opérations linéaries*, Warsaw 1932, p. 41, Theorem 5.

functions makes the group complete, metric, and separable. If (x, y) and $(x, y)_1$ are two distance functions defined on the same set of points then the statement that $(x_n, x) \to 0$ implies $(x_n, x_1) \to 0$ is equivalent to the statement that $\mathfrak{G}_1 \subset \mathfrak{G}$ where $\mathfrak{G}, \mathfrak{G}_1$ are the families of open sets determined by the distance functions (x, y), $(x, y)_1$ respectively. It is this latter form, i.e., $\mathfrak{G}_1 \subset \mathfrak{G}$, that we prefer to use in stating the assumption regarding the comparison of the two topologies. We shall state here two additional assumptions either one of which when combined with $\mathfrak{G}_1 \subset \mathfrak{G}$ will yield the desired conclusion, i.e., $\mathfrak{G}_1 = \mathfrak{G}$.

(H_1) X is a *locally complete and locally separable metric space* and $\mathfrak{N}$ is the family of open spheres in X. The family $\mathfrak{N}_1$ is a family of subsets of X forming a Hausdorff space (i.e., $\mathfrak{N}_1$ satisfies (i), (ii)). *For every $N \in \mathfrak{N}$ we have N not non-dense in the $\mathfrak{N}_1$ topology.*

(H_2) In this assumption X remains general (i.e., is not necessarily metric) and $\mathfrak{N}_1$ is as in H_1. The assumption H_2 is then $\mathfrak{N} \subset \sum_\phi \mathfrak{B}^1_\phi II^1_\phi$. The families $\mathfrak{B}^1_\phi$, II^1_ϕ are the Baire sets and second category sets resp. (with respect to the cardinal ϕ) in the $\mathfrak{N}_1$ topology.

In §5 it is shown that H_1 implies H_2. It is obvious that H_2 does not imply H_1 so that H_2 is the more general assumption but has the disadvantage that it is not as easily applied as H_1. *In case X is a complete metric and separable group and a new topology $\mathfrak{N}_1 \subset \mathfrak{G}$ is introduced which makes X a Hausdorff group of second category in itself then it is clear that H_1 and H_2 are both satisfied.*

Theorem 6.1: *Suppose that a new topology defined by a new neighborhood family $\mathfrak{N}_1 \subset \mathfrak{G}$ has been introduced in X forming a Hausdorff space X_1 which is also locally homogeneous with the same family $\mathfrak{H}$. Then either H_1 or H_2 will imply the equality $\mathfrak{G}_1 = \mathfrak{G}$.*

The identity map $x = f(x)$ on X to X_1 is $\mathfrak{H}$-distributive, for if $N_1 \in \mathfrak{N}_1$ then $h(N_1) \in \mathfrak{G}_1$. Furthermore since $\mathfrak{N}_1 \subset \mathfrak{G}$ the function f is continuous. Since f is 1–1 we have, by Lemma 1 above and Theorems 5.1 and 5.5, an interior point in $f(N)$ for every N in $\mathfrak{N}$. Thus by Lemma 4, $N = f(N) \in \mathfrak{G}_1$ and so $\mathfrak{G} \subset \mathfrak{G}_1 \subset \mathfrak{G}$.

6.3. *Additive mappings between groups.* A mapping $x_1 = f(x)$ of one group X onto another X_1 (both groups written additively) is called *additive* if $f(x + y) = f(x) + f(y)$. Such a function is necessarily $\mathfrak{H}$-distributive in the sense used above.

Theorem 6.2: *Let X_1 be a Hausdorff group and X a locally complete and locally separable metric group with a one-sided invariant metric.*[16] *Let f be a continuous additive function with domain X and range X_1 which takes spheres into sets which are not non-dense.*[17] *Then f is an interior transformation, i.e., f takes open sets into open sets.*

---

[16] For example $(x - y, 0) = (x, y)$. In view of the result of Kakutani, Proc. Imp. Acad. Jap., 12(1936), pp. 82–84, this hypothesis may be dispensed with. See also G. Birkhoff, Compositio Math., 3(1936), pp. 427–430. Birkhoff does not state explicitly the invariant property of the metric but it can be easily obtained from his work.

[17] This property is automatically satisfied if X is separable and X_1 is of second category in itself.

The one-sided invariant metric shows us that f is distributive in the sense in which that term is used in §5. The present theorem then follows immediately from Theorem 5.5 together with Lemma 4 above.

Finally it might be mentioned that Lemma 4 enables one to state immediately (i.e., without recourse to §5) a theorem of the above type for locally compact groups in which case the whole situation is much simpler.

YALE UNIVERSITY,
NEW HAVEN, CONN.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

ON CONTINUOUS MAPPINGS OF MANIFOLDS INTO SPHERES

BY SAMUEL EILENBERG

(Received June 19, 1939)

Introduction[1]

I. We shall consider in this paper continuous mappings of an $(n + 1)$-dimensional finite and oriented manifold[2] M^{n+1} into the $(m + 1)$-dimensional spherical manifold S^{m+1}.

Let p and q be the poles of S^{m+1}. A map $f(M^{n+1}) \subset S^{m+1}$ will be called *regular* if the sets $f^{-1}(p)$ and $f^{-1}(q)$ are $(n - m)$-dimensional polyhedra.[3]

(A) *Each mapping* $f(M^{n+1}) \subset S^{m+1}$ *is homotopic to a regular mapping* $g(M^{n+1}) \subset S^{m+1}$.

For each regular map $f(M^{n+1}) \subset S^{m+1}$ two $(n - m)$-dimensional cycles with integer coefficients: $\gamma_p^{n-m}(f)$ in $f^{-1}(p)$ and $\gamma_q^{n-m}(f)$ in $f^{-1}(q)$, will be defined so as to satisfy the following conditions:

(B) $$\gamma_p^{n-m}(f) \sim \gamma_q^{n-m}(f).$$

(C) *If g is a regular map homotopic to f, then* $\gamma_p^{n-m}(g) \sim \gamma_p^{n-m}(f)$.

(D) *Given a cycle* γ^{n-m} *such that* $\gamma^{n-m} \sim \gamma_p^{n-m}(f)$ *in* $M^{n+1} - f^{-1}(q)$, *there is an* $(n - m - 1)$*-dimensional subpolyhedron* P^{n-m-1} *of* M^{n+1} *and a map g homotopic to f such that*[4]

$$g^{-1}(p) = |\gamma^{n-m}| + P^{n-m-1}, \quad \gamma_p^{n-m}(g) = \gamma^{n-m},$$
$$g^{-1}(q) = f^{-1}(q), \qquad \gamma_q^{n-m}(g) = \gamma_q^{n-m}(f).$$

Denoting by $\Gamma(f)$ the element of the $(n - m)$-th homology group of M^{n+1} determined by $\gamma_p^{n-m}(f)$, we deduce from (A) and (C) that $\Gamma(f)$ is a homotopy invariant which may be regarded as defined for any map $f(M^{n+1}) \subset S^{m+1}$ not necessarily regular.

II. Let $n = 2m$, and let $f(M^{2m+1}) \subset S^{m+1}$ be a regular map such that $\Gamma(f) = 0$. We have then $\gamma_p^{n-m}(f) \sim 0$ and by (B) also $\gamma_q^{n-m}(f) \sim 0$. The linkage coefficient

$$c(f) = v[\gamma_p^{n-m}(f), \gamma_q^{n-m}(f)]$$

is therefore well determined.

---

[1] The paper is arranged so that all the main results are stated in the introduction which is divided into five parts. In five sections which follow the introduction, the corresponding proofs are given.

[2] See K. REIDEMEISTER, *Topologie de Polyeder*, Leipzig 1938, p. 151.

[3] With respect to a certain subdivision of M^{n+1}. The term *k-dimensional* when applied to subpolyhedra of M^{n+1} will always mean *at most k-dimensional*.

[4] $|\gamma^{n-m}|$ is the smallest $(n - m)$-dimensional polyhedron containing γ^{n-m}.

(E) *If g is a regular map homotopic to f, then* $c(g) = c(f)$.

(F) *If m is even,* $c(f) = 0$.

It follows from (A), (C) and (E) that $c(f)$ may be regarded as a homotopy invariant defined for any map $f(M^{2m+1}) \subset S^{m+1}$ such that $\Gamma(f) = 0$.

The invariants $\Gamma(f)$ and $c(f)$ have been introduced and investigated by Hopf.[5]

III. The property (D) announced above is essentially new and will make available a closer discussion of the mappings $f(M^{n+1}) \subset S^{m+1}$ for $m < n \leqq 2m$.

THEOREM I. *If* $m < n < 2m$ *then for a given mapping* $f(M^{n+1}) \subset S^{m+1}$ *the following properties are equivalent*:

(i) $\Gamma(f) = 0$;

(ii) *there is a mapping g homotopic to f such that* $g^{-1}(p)$ *and* $g^{-1}(q)$ *are* $(n - m - 1)$*-dimensional polyhedra*;

(iii) *there is a mapping g homotopic to f such that* $g^{-1}(p)$ *is a* $(n - m - 1)$*-dimensional polyhedron.*

THEOREM II. *Given a mapping* $f(M^{2m+1}) \subset S^{m+1} (m > 0)$ *the following properties are equivalent*:

(i) $c(f) = 0$;[6]

(ii) *there is a mapping g homotopic to f such that* $g^{-1}(p)$ *and* $g^{-1}(q)$ *are* $(m - 1)$*-dimensional polyhedra*;

(iii) *there is a mapping g homotopic to f such that* $g^{-1}(p)$ *is a* $(m - 1)$*-dimensional polyhedron.*

Let a^{2m+1} be a simplex of M^{2m+1} and let C_1^{m+1} and C_2^{m+1} be two oriented convex $(m + 1)$-dimensional simplicial cells inside a^{2m+1}. The oriented boundary spheres will be denoted by s_1^m and s_2^m, the corresponding cycles by σ_1^m and σ_2^m. If $s_1^m \cdot s_2^m = 0$ and if $v[\sigma_1^m, \sigma_2^m] = 1$, we shall say that s_1^m and s_2^m form an *elementary couple.*

THEOREM III. *Consider an elementary couple* s_1^m, s_2^m *in* $M^{2m+1}(m > 0)$ *and two integers* $\alpha_1 \neq 0$, $\alpha_2 \neq 0$. *Given a mapping* $f(M^{2m+1}) \subset S^{m+1}$ *we have* $c(f) = \alpha_1\alpha_2$,[6] *if, and only if, there are two* $(m - 1)$*-dimensional polyhedra* P^{m-1} *and* Q^{m-1} *and a mapping g homotopic to f such that*

$$g^{-1}(p) = s_1^m + P^{m-1}, \qquad \gamma_p^m(g) = \alpha_1\sigma_1^m,$$
$$g^{-1}(q) = s_2^m + Q^{m-1}, \qquad \gamma_q^m(g) = \alpha_2\sigma_2^m.$$

IV. The conditions in Th. I–III can be simplified if M^{n+1} is subject to the following homology condition:[7]

(H) $\quad B^i(M^{n+1})$ coef $(m^{n-i}) = 0$ for $i = 1, 2, \cdots, n - m - 1$.

In the case $n = 2m$ condition (H) takes the following form:

(H′) $\quad B^i(M^{2m+1})$ coef $(m^{2m-i}) = 0$ for $i = 1, 2, \cdots, m - 1$.

[5] H. HOPF, *Math. Ann.* 104 (1931), pp. 637–665, and *Fund. Math.* 25 (1935), pp. 427–440.

[6] This includes $\Gamma(f) = 0$.

[7] (m^k) stands for the k-th homotopy group of S^m. The group is abelian and therefore can be taken as a coefficient group for $B^i(M^{n+1})$. See W. HUREWICZ, *Proc. Akad. Amsterdam* 38 (1935), p. 113.

Assuming that M^{n+1} satisfies (H) we restate our theorems as follows:

THEOREM Ia. *If $m < n < 2m$ then for a given mapping $f(M^{n+1}) \subset S^{m+1}$ the following properties are equivalent*:

(i) $\Gamma(f) = 0$;

(ii) *there is a mapping g homotopic to f such that $g^{-1}(p)$ and $g^{-1}(q)$ are single points*;

(iii) *there is a mapping g homotopic to f such that $g^{-1}(p)$ is a single point.*

THEOREM IIa. *Given a mapping $f(M^{2m+1}) \subset S^{m+1} (m > 0)$ the following properties are equivalent*:

(i) $c(f) = 0$;[6]

(ii) *there is a mapping g homotopic to f such that $g^{-1}(p)$ and $g^{-1}(q)$ are single points*;

(iii) *there is a mapping g homotopic to f such that $g^{-1}(p)$ is a single point.*

THEOREM IIIa. *Consider an elementary couple s_1^m, s_2^m in M^{2m+1} $(m > 0)$ and two integers $\alpha_1 \neq 0$, $\alpha_2 \neq 0$. Given a mapping $f(M^{2m+1}) \subset S^{m+1}$ we have $c(f) = \alpha_1\alpha_2$[6] if and only if there is a mapping g homotopic to f such that*

$$g^{-1}(p) = s_1^m, \qquad \gamma_p^m(g) = \alpha_1\sigma_1^m,$$

$$g^{-1}(q) = s_2^m, \qquad \gamma_q^m(g) = \alpha_2\sigma_2^m.$$

The condition (H) is always satisfied if $m = n + 1$. It follows that *Th. Ia can always be applied to mappings $f(M^{m+2}) \subset S^{m+1} (m > 1)$* and that *Th. IIa and IIIa hold for mappings $f(M^3) \subset S^2$.*

The groups (m^k) for $k > m$ are generally unknown. They have been calculated[8] in the cases $k = m + 1$ and $k = m + 2$.

Using these results we verify that:

Th. IIa can be applied to mappings $f(M^5) \subset S^3$ provided $B^1(M^5) = 0$ (coef. integers).[9]

Th. Ia can be applied to mappings $f(M^{m+3}) \subset S^{m+1} (m \geqq 3)$ provided $B^1(M^{m+3}) = 0$ (coef. mod 2).

Th. IIa and IIIa can be applied to mappings $f(M^7) \subset S^4$ provided $B^2(M^7) = 0$ (coef. mod 2).

Th. Ia can be applied to mappings $f(M^{m+4}) \subset S^{m+1} (m \geqq 4)$ provided $B^2(M^{m+4}) = 0$ (coef. mod 2).

V. The most important case when (H) is always satisfied is when $M^{n+1} = S^{n+1}$. It follows that Th. Ia–IIIa can then always be applied. Using certain operations introduced by Freudenthal[10] and Hopf[11] I shall be able to restate these theorems in a more geometrical way.

Let $S^{m+1} = E_+^{m+1} + E_-^{m+1}$ be a decomposition of S^{m+1} into two hemispheres such that: $S^m = E_+^{m+1} \cdot E_-^{m+1}$ is the equator of S^{m+1}, p is the "center" of E_+^{m+1}

[8] L. PONTRJAGIN, *C. R. Acad. Sci. URSS* 19 (1938), pp. 147–149, 361–363.

[9] Theorem IIIa has no application in this case because of (F).

[10] H. FREUDENTHAL, *Compositio Math.* 5 (1937), pp. 299–314.

[11] H. HOPF, *Fund. Math.* 25 (1935), pp. 427–440.

and q is the center of E_-^{m+1}. Similarly we have a division $S^{n+1} = E_+^{n+1} + E_-^{n+1}$ such that $S^n = E_+^{n+1} \cdot E_-^{n+1}$. Let p' and q' be the "centers" of E_+^{n+1} and E_-^{n+1}.

Given a mapping $\varphi(S^n) \subset S^m$ we define a mapping $F^\varphi(S^{n+1}) \subset S^{m+1}$ by the following condition: for each point $x \in S^n$ the radius p', x of E_+^{n+1} is mapped linearly onto the radius $p, \varphi(x)$ of E_+^{m+1} and the radius q', x of E_-^{n+1} is mapped linearly onto the radius $q, \varphi(x)$ of E_-^{m+1}.

The operation F^φ was established by Freudenthal[10] who also proved the following two theorems:

THEOREM Ib. *If $m < n < 2m$ then for every mapping $f(S^{n+1}) \subset S^{m+1}$ there is a mapping $\varphi(S^n) \subset S^m$ such that F^φ is homotopic to f.*

THEOREM IIb. *Given a mapping $f(S^{2m+1}) \subset S^{m+1}(m > 0)$ we have $c(f) = 0$ if and only if there is a mapping $\varphi(S^{2m}) \subset S^m$ such that F^φ is homotopic to f.*

Note that according to (F) we have $c(f) = 0$ if m is even.

A mapping $\varphi(S^m \times S^m) \subset S^m$ of the topological product $S^m \times S^m$ is said[12] to be of the type (α_1, α_2) if for a given $t \in S^m$, $S^m \times t$ is mapped with the degree α_1 and $t \times S^m$ is mapped with the degree α_2.

Let E^{m+1} be the full sphere bounded by S^m, and let σ be the center of E^{m+1}. The boundary $E^{m+1} \times S^m + S^m \times E^{m+1}$ of the product $E^{m+1} \times E^{m+1}$ may be regarded as identical with S^{2m+1}. Given a map $\varphi(S^m \times S^m) \subset S^m$ we define a map $H^\varphi(E^{m+1} \times S^m + S^m \times E^{m+1}) \subset S^{m+1}$ by the following condition: for each point $(x_1, x_2) \in S^m \times S^m$ the radius $(0, x_2), (x_1, x_2)$ of $E^{m+1} \times x_2$ is mapped linearly onto the radius $p, \varphi(x_1, x_2)$ of E_+^{m+1} and the radius $(x_1, 0), (x_1, x_2)$ of $x_1 \times E^{m+1}$ is mapped linearly onto the radius $q, \varphi(x_1, x_2)$ of E_-^{m+1}.

The operation H^φ has been introduced by Hopf[11] who also proved the theorem that if φ is of the type (α_1, α_2) then $c(H^\varphi) = \alpha_1\alpha_2$. The following theorem may be considered as its inverse.

THEOREM IIIb. *Given a mapping $f(S^{2m+1}) \subset S^{m+1}(m > 0)$ such that $c(f) = \alpha_1 \cdot \alpha_2 \neq 0$ there is a mapping $\varphi(S^m \times S^m) \subset S^m$ of the type (α_1, α_2) such that H^φ is homotopic to f.*

For even m it was proved by Hopf[13] that each mapping $\varphi(S^m \times S^m) \subset S^m$ is either of the type $(\alpha_1, 0)$ or of the type $(0, \alpha_2)$.

On the other hand Freudenthal[14] has recently announced a theorem that if m is odd and k is an integer, there is a map $f(S^{2m+1}) \subset S^{m+1}$ with $c(f) = k$. Using Th. IIIa we obtain that

If m is odd and α_1, α_2 are any two integers, there is a mapping $\varphi(S^m \times S^m) \subset S^m$ of the type (α_1, α_2).

This solves a problem proposed by Hopf.[15]

I. Definition and Properties of $\gamma_p^{n-m}(f)$

1. *Proof of* (A). It is well known that there is a simplicial map $g(M^{n+1}) \subset S^{m+1}$ homotopic to f. The simplicial division of S^{m+1} may be chosen

[12] H. HOPF, loc cit., p. 431.
[13] Ibid., p. 436. See also H. WHITNEY, *Ann. of Math.* 39 (1938), p. 428.
[14] H. FREUDENTHAL, *Proc. Akad. Amsterdam* 42 (1939), p. 140.
[15] Ibid., p. 436.

so that p and q are inside points of some $(m+1)$-dimensional simplexes. It is easy to verify that g is regular.

2. Let $f(M^{n+1}) \subset S^{m+1}$ be a regular map. We consider M^{n+1} in a subdivision such that $f^{-1}(p)$ and $f^{-1}(q)$ are subcomplexes. To this simplicial division there corresponds a dual cell division so that to each k-simplex a^k in M^{n+1} there corresponds a $(n+1-k)$-cell b^{n+1-k}.

Let G be an arbitrary abelian group. Given a k-chain $A^k = \sum \alpha_i a_i^k$ with integer coefficients, and a $(n+1-k)$-chain $B^{n+1-k} = \sum \beta_i b_i^{n+1-k}$ with coefficients in G, we shall consider the intersection coefficient (which is an element of G) $\chi[A^k, B^{n+1-k}] = \sum \alpha_i \beta_i$.

For each $(m+1)$-dimensional chain $B^{m+1} = \sum \beta_i b_i^{m+1}$ we have $\partial B^{m+1} \subset M^{n+1} - f^{-1}(p)$ and therefore $f(\partial B^{m+1}) \subset S^{m+1} - p$. It follows that the *local degree* in the point p, $\mathfrak{g}[f, B^{m+1}, p] \in G$, is defined. If B^{m+1} is a cycle then the *degree* $\mathfrak{g}[f, B^{m+1}]$ is defined and $\mathfrak{g}[f, B^{m+1}] = \mathfrak{g}[f, B^{m+1}, p]$.

3. *Definition of* $\gamma_p^{n-m}(f)$. Let a_i^{n-m} be a simplex of $f^{-1}(p)$. We consider the integer $\alpha_i = \mathfrak{g}[f, b_i^{m+1}, p]$ and the $(n-m)$-chain $\gamma_p^{n-m}(f) = \sum \alpha_i a_i^{n-m}$ in $f^{-1}(p)$. Since $\chi[\gamma_p^{n-m}(f), b_i^{m+1}] = \alpha_i = \mathfrak{g}[f, b_i^{m+1}, p]$ for each dual cell b_i^{m+1}, therefore

(1) $\quad \chi[\gamma_p^{n-m}(f), B^{m+1}] = \mathfrak{g}[f, B^{m+1}, p]$ *for each chain* $B^{m+1} = \sum \beta_i b_i^{m+1}$,

and

(2) $\quad \chi[\gamma_p^{n-m}(f), B^{m+1}] = \mathfrak{g}[f, B^{m+1}]$ *if* B^{m+1} *is a cycle.*

Let b^{m+2} be a dual $(m+2)$-cell. ∂b^{m+2} is then a $(m+1)$-cycle and $\mathfrak{g}[f, \partial b^{m+2}] = 0$ since $\partial b^{m+2} \sim 0$. By (2) we have then $\chi[\gamma_p^{n-m}(f), \partial b^{m+2}] = 0$ and therefore $\chi[\partial \gamma_p^{n-m}(f), b^{m+2}] = 0$. Since this holds for each $(m+2)$-cell b^{m+2} therefore $\partial \gamma_p^{n-m}(f) = 0$ and $\gamma_p^{n-m}(f)$ is a $(n-m)$-cycle.

Similarly the cycle $\gamma_q^{n-m}(f)$ in $f^{-1}(q)$ is defined, and (1) and (2) hold.

4. *Proof of* (B). Applying (2) to $\gamma_p^{n-m}(f)$ and $\gamma_q^{n-m}(f)$ we find that $\chi[\gamma_p^{n-m}(f), B^{m+1}] = \chi[\gamma_q^{n-m}(f), B^{m+1}]$ for each cycle B^{m+1}. This implies[16] $\gamma_p^{n-m}(f) \sim \gamma_q^{n-m}(f)$.

5. *Proof of* (C). For each cycle B^{m+1} we have $\mathfrak{g}[g, B^{m+1}] = \mathfrak{g}[f, B^{m+1}]$. Therefore by (2) $\chi[\gamma_p^{n-m}(g), B^{m+1}] = \chi[\gamma_p^{n-m}(f), B^{m+1}]$. This implies[16] $\gamma_p^{n-m}(g) \sim \gamma_p^{n-m}(f)$.

6. We consider S^{m+1} as the set of points $x = (x_1, x_2, \cdots, x_{m+1}, x_{m+2})$ of the euclidean $(m+2)$-space, defined by the condition $x_1^2 + x_2^2 + \cdots + x_{m+1}^2 + x_{m+2}^2 = 1$. Let us put $l(x) = x_{m+2}$ for $x \in S^{m+1}$. The point p is then defined by the condition: $l(p) = 1$, the point q by $l(q) = -1$, the equator S^m by the condition $l(x) = 0$.

[16] By Poincaré's duality theorem.

For each point $x \in S^{m+1} - p - q$ we define

$$r(x) = \left(\frac{x_1}{(1 - x_{m+2}^2)^{1/2}}, \frac{x_2}{(1 - x_{m+2}^2)^{1/2}}, \dots, \frac{x_{m+1}}{(1 - x_{m+2}^2)^{1/2}}, 0\right).$$

This way we obtain a mapping $r(S^{m+1} - p - q) \subset S^m$ which is a "radial" projection.

Given a point $y \in S^m$ and a number $-1 < l < 1$ there is a point $x = \{y, l\} \in S^{m+1} - p - q$ uniquely defined, such that $r(x) = y$, $l(x) = l$.

7. *Proof of* (D). Let V be the subcomplex of M^{n+1} consisting of all the simplices of M^{n+1} which have no vertex on $f^{-1}(q)$. Let U be an open set such that $f^{-1}(q) \subset U \subset \bar{U} \subset M^{n+1} - V$. Let $P^{n-m} = f^{-1}(p)$, $X = M^{n+1} - U$, $Z = \bar{U} \cdot X$, and $\varphi(x) = r[f(x)]$ for $x \in X - P^{n-m}$. We obtain this way a map $\varphi(X - P^{n-m}) \subset S^m$.

Let a_i^{n-m} be a $(n - m)$-simplex of P^{n-m} and b_i^{m+1} its dual cell. Since $b_i^{m+1} \subset V \subset X - Z$ therefore $\mathfrak{g}[\varphi, \partial b_i^{m+1}] = \mathfrak{g}[f, b_i^{m+1}, p]$ and according to the definition of $\gamma_p^{n-m}(f)$ we then have $\gamma_p^{n-m}(f) = \sum \mathfrak{g}[\varphi, \partial b_i^{m+1}]a_i^{n-m}$. It follows that $\gamma_p^{n-m}(f) = \gamma^{n-m}(\varphi)$ where $\gamma^{n-m}(\varphi)$ is a $(n - m)$-cycle attached to the map $\varphi(X - P^{n-m}) \subset S^m$ which I have defined in another paper.[17]

The relation $\gamma^{n-m} \sim \gamma_p^{n-m}(f)$ in $M^{n+1} - f^{-1}(q)$ implies $\gamma^{n-m} \sim \gamma_p^{n-m}(f)$ in $V \subset X - Z$ and therefore $\gamma^{n-m} \sim \gamma^{n-m}(\varphi)$ in $X - Z$. Therefore by a lemma[18] concerning the cycle $\gamma^{n-m}(\varphi)$ there is a $(n - m - 1)$-dimensional polyhedron $P^{n-m-1} \subset X - Z$ and a map $\psi(X - |\gamma^{n-m}| - P^{n-m-1}) \subset S^m$ such that $\gamma^{n-m}(\psi) = \gamma^{n-m}$, $\psi(x) = \varphi(x)$ if $x \in Z$.

Let $k(x)$ be a continuous function defined for $x \in M^{n+1}$, $-1 \leqq k(x) \leqq 1$, and such that

$$k(x) = -1 \quad \text{if and only if} \quad x \in f^{-1}(q),$$
$$k(x) = 1 \quad \text{if and only if} \quad x \in |\gamma^{n-m}| + P^{n-m-1},$$
$$k(x) = l[f(x)] \quad \text{if} \quad x \in U.$$

We define the map $g(M^{n+1}) \subset S^{m+1}$ by taking

$$g(x) = f(x) \quad \text{for} \quad x \in U,$$
$$g(x) = \{\psi(x), k(x)\} \quad \text{for} \quad x \in M^{n+1} - U - |\gamma^{n-m}| - P^{n-m-1},$$
$$g(x) = p \quad \text{for} \quad x \in |\gamma^{n-m}| + P^{n-m-1}.$$

The condition $g^{-1}(p) = |\gamma^{n-m}| + P^{n-m-1}$ and $g^{-1}(q) = f^{-1}(q)$ are obviously satisfied. Since $r[g(x)] = \psi(x)$ for each $x \in X - g^{-1}(p)$ therefore $\gamma_p^{n-m}(g) = \gamma^{n-m}(\psi) = \gamma^{n-m}$. Since $g(x) = f(x)$ for $x \in U$ and $g^{-1}(q) = f^{-1}(q) \subset U$ therefore $\gamma_q^{n-m}(g) = \gamma_q^{n-m}(f)$ and there is a neighborhood U_1 of q in S^{m+1} such that

[17] *Fund. Math.* 31 (1938), p. 183.
[18] Ibid., p. 198.

$g^{-1}(x) = f^{-1}(x)$ if $x \epsilon U_1$. The existence of such an U_1 implies easily that g and f are homotopic.

Note that (C) and (D) can be restated replacing p by q and vice versa.

8. It follows from Poincare's duality theorem that the homology class $\Gamma(f)$ of $\gamma_p^{n-m}(f)$ is defined by the property (2). In particular $\Gamma(f) = 0$ is characterized by the condition $\mathfrak{g}[f, B^{m+1}] = 0$ for each cycle B^{m+1}. This means that $\Gamma(f) = 0$ if and only if the map $f(M^{n+1}) \subset S^{m+1}$ is algebraically inessential.

If $n = m$ then $\gamma_p^{n-m}(f)$ is a 0-dimensional cycle of $\Gamma(f)$ which can be considered as an integer. It follows from (2) that $\Gamma(f) = \mathfrak{g}[f, M^{n+1}]$.

II. Properties of $c(f)$

1. *Proof of* (E). Let C be the subset of S^{m+1} defined by the condition $-\frac{1}{2} \leqq l(x) \leqq \frac{1}{2}$. Let $\eta > 0$ be a number having the following property: if $x \epsilon C$, $y \epsilon S^{m+1}$ and $| x - y | < \eta$ then $y \epsilon S^{m+1} - p - q$ and $| r(x) - r(y) | < 2$.

Given two regular and homotopic maps f and g there is a finite sequence of maps $f_0 = f, f_1, f_2, \cdots, f_k = g$ such that $| f_i(x) - f_{i+1}(x) | < \eta$ if $x \epsilon M^{2m+1}$ and $i = 0, 1, \cdots, k - 1$. By (A) we may admit that all these maps are regular. The proof of (E) reduces therefore to the case when

$$| f(x) - g(x) | < \eta \quad \text{for} \quad x \epsilon M^{2m+1}.$$

Let $M^{2m+1} = R_1 + R_2$ be a decomposition of M^{2m+1} into two polyhedra such that $f^{-1}(p) \subset R_1 - R_2$, $f^{-1}(q) \subset R_2 - R_1$, $f(R_1 \cdot R_2) \subset C$. Writing $\varphi(x) = r[f(x)]$ and $\psi(x) = r[g(x)]$ for $x \epsilon R_1 \cdot R_2$, we have by definition of η, $| \varphi(x) - \psi(x) | < 2$ if $x \epsilon R_1 \cdot R_2$. It follows that φ and ψ are homotopic and therefore that $\mathfrak{g}[\varphi, B^m] = \mathfrak{g}[\psi, B^m]$ for each cycle $B^m \subset R_1 \cdot R_2$. In particular if B^{m+1} is a chain in R_1 such that $\partial B^{m+1} \subset R_1 \cdot R_2$ we have $\mathfrak{g}[\varphi, \partial B^{m+1}] = \mathfrak{g}[\psi, \partial B^{m+1}]$. This implies $\mathfrak{g}[f, B^{m+1}, p] = \mathfrak{g}[g, B^{m+1}, p]$ and by (1) $\chi[\gamma_p^m(f), B^{m+1}] = \chi[\gamma_p^m(g), B^{m+1}]$. By duality theorems[19] we then have $\gamma_p^m(f) \sim \gamma_p^m(g)$ in $R_1 - R_2$. Similarly $\gamma_q^m(f) \sim \gamma_q^m(g)$ in $R_2 - R_1$ and therefore $c(f) = c(g)$.

2. *Proof of* (F). Let $h(S^{m+1}) = S^{m+1}$ be a homeomorphic map of degree 1 such that $h(p) = q$ and $h(q) = p$. Let $g(x) = h[f(x)]$. As it easily follows from the definition, we have $\gamma_p^m(g) = \gamma_q^m(f)$, $\gamma_q^m(g) = \gamma_p^m(f)$. Since f and g are homotopic we have $c(f) = c(g)$, and therefore $c(f) = v[\gamma_p^m(f), \gamma_q^m(f)] = v[\gamma_q^m(f), \gamma_p^m(f)]$. However, if m is even $v[\gamma_p^m(f), \gamma_q^m(f)] = -v[\gamma_q^m(f), \gamma_p^m(f)]$ and thus $c(f) = 0$.

III. The Case $m < n \leqq 2m$

1. We shall often use the following elementary

LEMMA 1. *Let R be a subpolyhedron of M^{n+1} and R^{k_1} a k_1-dimensional subpolyhedron of M^{n+1}. Given a k_2-cycle $\gamma^{k_2} \subset M^{n+1} - R - R^{k_1}$, $k_1 + k_2 < n$, then $\gamma^{k_2} \sim 0$ in $M^{n+1} - R$ implies $\gamma^{k_2} \sim 0$ in $M^{n+1} - R - R^{k_1}$.*

[19] See L. PONTRJAGIN, *Math. Ann.* 105 (1931), p. 190; also *Ann. of Math.* 35 (1934), p. 312.

2. We shall consider m-cycles γ^m in M^{2m+1} having the following property (b) *each m-simplex of $|\gamma^m|$ has in γ^m a coefficient ± 1; for each cycle γ_1^m in $|\gamma^m|$ with coefficients in a group G there is an element $\alpha \in G$ such that $\gamma_1^m = \alpha\gamma^m$.*

Using duality theorems[19] we obtain the following

LEMMA 2. *Let $\gamma^m \subset M^{2m+1}$ be a cycle having the property* (b) *and such that $\gamma^m \sim 0$. Given a cycle $\gamma_1^m \subset M^{2m+1} - |\gamma^m|$ (coefficients in G) such that $\gamma_1^m \sim 0$, we have $\gamma_1^m \sim 0$ in $M^{2m+1} - |\gamma^m|$ if and only if $v[\gamma^m, \gamma_1^m] = 0$.*

3. LEMMA 3. *Given a mapping $f(M^{2m+1}) \subset S^{m+1}$ such that $\Gamma(f) = 0$, there is a $(m-1)$-dimensional subpolyhedron P^{m-1} of M^{2m+1} and a regular mapping g homotopic to f such that $\gamma_p^m(g)$ has the property* (b) *and $g^{-1}(p) = |\gamma_p^m(g)| + P^{m-1}$.*

PROOF. Let g be a normal map homotopic to f. Since $\Gamma(g) = \Gamma(f) = 0$ therefore $\gamma_p^m(g) \sim 0$. By a general theorem[20] there is a cycle $\gamma^m \subset M^{2m+1} - g^{-1}(q)$ having the property (b) and such that $\gamma^m \sim \gamma_p^m(g)$ in $M^{2m+1} - g^{-1}(q)$. Using (D) we obtain a polyhedron P^{m-1} and a regular map g_1 homotopic to g such that $g_1^{-1}(p) = |\gamma^m| + P^{m-1}$, $\gamma_p^m(g_1) = \gamma^m$.

4. *Proof of* Th. I *and* II, (i) $\Longrightarrow$ (ii). We first find a regular map g homotopic to f such that $\gamma_q^{n-m}(g) \sim 0$ in $M^{n+1} - g^{-1}(p)$.

If $n < 2m$ we may take g to be any regular map homotopic to f. Since $\Gamma(g) = \Gamma(f) = 0$ therefore $\gamma_q^{n-m}(g) \sim 0$ and by Lemma 1 we have also $\gamma_q^{n-m}(g) \sim 0$ in $M^{n+1} - g^{-1}(p)$.

If $n = 2m$ we choose g so as to satisfy Lemma 3. Since $c(g) = c(f) = 0$ therefore $v[\gamma_p^m(g), \gamma_q^m(g)] = 0$. The cycle $\gamma_p^m(g)$ satisfying the condition (b) it follows from Lemma 2 that $\gamma_q^m(g) \sim 0$ in $M^{2m+1} - |\gamma_p^m(g)|$. Applying Lemma 1 we obtain $\gamma_q^m(g) \sim 0$ in $M^{2m+1} - |\gamma_p^m(g)| - P^{m-1} = M^{2m+1} - g^{-1}(p)$.

The existence of g being thus proved, there is by (D) a $(n-m-1)$-dimensional polyhedron Q^{n-m-1} and a regular map g_1 homotopic to g such that $g_1^{-1}(q) = Q^{n-m-1}$.

Since $\Gamma(g_1) = 0$ we have $\gamma_p^{n-m}(g_1) \sim 0$ and by Lemma 1 also $\gamma_p^{n-m}(g_1) \sim 0$ in $M^{n+1} - Q^{n-m-1}$. Using (D) once more we obtain a $(n-m-1)$-dimensional polyhedron P^{n-m-1} and a map g_2 homotopic to g_1 such that $g_2^{-1}(p) = P^{n-m-1}$, $g_2^{-1}(q) = Q^{n-m-1}$.

(ii) $\Longrightarrow$ (iii). Clear.

(iii) $\Longrightarrow$ (i). Without restricting the generality we may admit that the map g given by (iii) is regular. We then have $\gamma_p^{n-m}(g) = 0$ and therefore $\Gamma(g) = 0$ (and $c(f) = 0$ if $n = 2m$).

5. *Proof of* Th. III. *Necessity.* Let $\Gamma(f) = 0$ and $c(f) = \alpha_1\alpha_2$. Because of the homogeneity of M^{2m+1} it is enough to prove that the condition of Th. III is satisfied for some elementary couple s_1^m, s_2^m which we shall construct during the proof.

---

[20] S. EILENBERG, See abstract No. 267, *Bull. A. M. S.* 46 (1940), p. 416. The paper in full was to have appeared in *Fund. Math.* 33 (1939).

Let g be the map given by Lemma 3, and let γ^m be a cycle in $M^{2m+1} - g^{-1}(p)$ such that $\gamma^m \sim 0$ in M^{2m+1} and that $v[\gamma_p^m(g), \gamma^m] = \alpha_1$. By a general theorem[20] there is a cycle γ_1^m having the property (b) and such that $\gamma_1^m \sim \gamma^m$ in $M^{2m+1} - g^{-1}(p)$. Since $v[\gamma_p^m(g), \gamma_1^m] = \alpha_1$ and $v[\gamma_p^m(g), \gamma_q^m(g)] = \alpha_1\alpha_2$ therefore according to Lemma 2 we have $\alpha_2\gamma_1^m \sim \gamma_q^m(g)$ in $M^{2m+1} - |\gamma_p^m(g)|$. Using Lemma 1 we therefore obtain $\alpha_2\gamma_1^m \sim \gamma_q^m(g)$ in $M^{2m+1} - g^{-1}(p)$. Applying (D) we obtain a $(m-1)$-dimensional polyhedron Q_1^{m-1} and a regular map g_1 homotopic to g such that $g_1^{-1}(q) = |\gamma_1^m| + Q_1^{m-1}$, $\gamma_q^m(g_1) = \alpha_2\gamma_1^m$. Now, let a_1^m be a simplex of $|\gamma_1^m|$, b_1^{m+1} its dual cell, s_1^m the boundary of b_1^{m+1} and σ_1^m the cycle ∂b_1^{m+1}. We may admit that a_1^m is oriented so as to have $v[\sigma_1^m, \gamma_1^m] = 1$ and therefore $v[\alpha_1\sigma_1^m, \gamma_1^m] = \alpha_1$. On the other hand we have $c(g_1) = \alpha_1\alpha_2$ and therefore $\alpha_1\alpha_2 = v[\gamma_p^m(g_1), \gamma_q^m(g_1)] = v[\gamma_p^m(g_1), \alpha_2\gamma_1^m]$. It follows that $v[\gamma_p^m(g_1), \gamma_1^m] = \alpha_1$. By Lemma 2 we then have $\alpha_1\sigma_1^m \sim \gamma_p^m(g_1)$ in $M^{2m+1} - |\gamma_1^m|$ and by Lemma 1 also $\alpha_1\sigma_1^m \sim \gamma_p^m(g_1)$ in $M^{2m+1} - |\gamma_1^m| - Q_1^{m-1} = M^{2m+1} - g_1^{-1}(q)$. Applying (D) we obtain a $(m-1)$-dimensional polyhedron P^{m-1} and a regular map g_2 homotopic to g_1 such that $g_2^{-1}(p) = s_1^m + P^{m-1}$, $\gamma_p^m(g_2) = \alpha_1\sigma_1^m$. Let a_2^m be a simplex of s_1^m, b_2^{m+1} the dual cell of a_2^m, s_2^m the boundary of a_2^m and let $\sigma_2^m = \partial b_2^{m+1}$. Clearly $v[\sigma_1^m, \alpha_2\sigma_2^m] = \alpha_2$. On the other hand we have $c(g_2) = \alpha_1\alpha_2$ and therefore $\alpha_1\alpha_2 = v[\gamma_p^m(g_2), \gamma_q^m(g_2)] = v[\alpha_1\sigma_1^m, \gamma_q^m(g_2)]$. It follows that $v[\sigma_1^m, \gamma_q^m(g_2)] = \alpha_2$. Since σ_1^m has the property (b) and $|\sigma_1^m| = s_1^m$, therefore by Lemma 2 $\alpha_2\sigma_2^m \sim \gamma_q^m(g_2)$ in $M^{2m+1} - s_1^m$ and by Lemma 1 also $\alpha_2\sigma_2^m \sim \gamma_q^m(g_2)$ in $M^{2m+1} - s_1^m - P^{m-1} = M^{2m+1} - g_2^{-1}(p)$. Applying (D) we obtain a $(m-1)$-dimensional polyhedron Q^{m-1} and a map g_3 homotopic to g_2 such that

$$g_3^{-1}(p) = g_2^{-1}(p) = s_1^m + P^{m-1}, \qquad \gamma_p^m(g_3) = \gamma_p^m(g_2) = \alpha_1\sigma_1^m,$$
$$g_3^{-1}(q) = s_2^m + Q^{m-1}, \qquad \gamma_q^m(g_3) = \alpha_2\sigma_2^m.$$

This finishes the proof, since s_1^m and s_2^m obviously form an elementary couple.

Sufficiency. Since $\sigma_1^m \sim 0$ therefore $\gamma_p^m(g) \sim 0$ and $\Gamma(g) = 0$. Since $v[\sigma_1^m, \sigma_2^m] = 1$, therefore $c(g) = v[\alpha_1\sigma_1^m, \alpha_2\sigma_2^m] = \alpha_1\alpha_2$.

IV. The Condition (H)

1. LEMMA 4. *Given a regular mapping $f(M^{n+1}) \subset S^{m+1}$ such that $f^{-1}(p) = P^{n-m} - P^{n-m-1}$ where P^{n-m-1} is a $(n-m-1)$-dimensional polyhedron, P^{n-m} a $(n-m)$-dimensional polyhedron such that*

(H₁) $B^i[M^{n+1} - f^{-1}(q)] \bmod P^{n-m} \text{ coef } (m^{n-i-1}) = 0$

for $i = 0, 1, \cdots, n-m-1$,

there is a mapping g homotopic to f such that

$$g^{-1}(p) = P^{n-m}, \qquad g^{-1}(q) = f^{-1}(q), \qquad \gamma_q^{n-m}(g) = \gamma_q^{n-m}(f).$$

PROOF. Since $f^{-1}(q)$ is a polyhedron and since $M^{n+1} - f^{-1}(q)$ is connected, there is a neighborhood U of $f^{-1}(q)$ such that

$$B^i(M^{n+1} - \bar{U}) \bmod P^{n-m} \text{ coef } (m^{n-i-1}) = 0 \quad \text{for} \quad i = 0, 1, \cdots, n-m-1.$$

Let $X = M^{n+1} - U$, $Z = \overline{U} \cdot X$, and $\varphi(x) = r[f(x)]$ for $x \epsilon X - P^{n-m} - P^{n-m-1}$. We obtain this way a map $\varphi(X - P^{n-m} - P^{n-m-1}) \subset S^m$. Since $B^i(X - Z)$ mod P^{n-m} coef $(m^{n-i-1}) = 0$ for $i = 0, 1, \cdots, n - m - 1$ therefore[21] there is a map $\psi(X - P^{n-m}) \subset S^m$ such that $\psi(x) = \varphi(x)$ if $x \epsilon Z$.

Let $k(x)$ be a continuous function defined for $x \epsilon M^{n+1}$, $-1 \leqq k(x) \leqq 1$, and such that

$$k(x) = -1 \quad \text{if and only if} \quad x \epsilon f^{-1}(q),$$

$$k(x) = 1 \quad \text{if and only if} \quad x \epsilon P^{n-m},$$

$$k(x) = l[f(x)] \quad \text{if} \quad x \epsilon U.$$

We define the map $g(M^{n+1}) \subset S^{m+1}$ by taking

$$g(x) = f(x) \quad \text{for} \quad x \epsilon U,$$

$$g(x) = \{\psi(x), k(x)\} \quad \text{for} \quad x \epsilon M^{n+1} - U - P^{n-m},$$

$$g(x) = p \quad \text{for} \quad x \epsilon P^{n-m}.$$

Conditions $g^{-1}(p) = P^{n-m}$ and $g^{-1}(q) = f^{-1}(q)$ are obviously satisfied. Since $g(x) = f(x)$ for $x \epsilon U$ and $g^{-1}(q) = f^{-1}(q) \subset U$, therefore $\gamma_q^{n-m}(g) = \gamma_q^{n-m}(f)$ and there is a neighborhood U_1 of q in S^{m+1} such that $g^{-1}(x) = f^{-1}(x)$ if $x \epsilon U_1$. The existence of such an U_1 easily implies that g and f are homotopic.

2. Now let us admit that M^{n+1} satisfies (H) and that P^{n-m} is non-vacuous, connected, and such that

$$(\mathrm{H}_2) \quad B^{i-1}(P^{n-m}) \text{ coef } (m^{n-m-1}) = 0 \quad \text{for} \quad i = 1, 2, \cdots, n - m - 1.$$

From (H) using Lemma 1 we obtain $B^i[M^{n+1} - f^{-1}(q)]$ coef $(m^{n-m-1}) = 0$ for $i = 1, 2, \cdots, n - m - 1$. From this and from (H_2) it follows that condition (H_1) is satisfied.

3. *Proof of* Th. Ia *and* IIa. (i) $\Longrightarrow$ (ii). By Th. I and II we may admit that $f^{-1}(p)$ and $f^{-1}(q)$ are $(n - m - 1)$-dimensional polyhedra. Let P^{n-m} consist of an arbitrary point of $f^{-1}(p)$. Conditions (H_2) are then satisfied and therefore condition (H_1) of Lemma 4 is satisfied. There is therefore a map g homotopic to f such that $g^{-1}(p)$ is a point and $g^{-1}(q) = f^{-1}(q)$. Repeating the same argument with f replaced by g and with q replaced by p we find a map g_1 homotopic to g and such that $g_1^{-1}(p)$, $g_1^{-1}(q)$ are single points.

(ii) $\Longrightarrow$ (iii). Clear. (iii) $\Longrightarrow$ (i). Follows from Th. I and II.

4. *Proof of* Th. IIIa. *Necessity.* Let g be a map homotopic to f given by Th. III. Taking $P^m = s_1^m$ we see that conditions (H_2) are satisfied. Condition (H_1) is therefore satisfied too and Lemma 4 can be applied to g. We

[21] S. Eilenberg, *Fund. Math.* 31 (1938), p. 184.

obtain therefore a map g_1 homotopic to g such that $g_1^{-1}(p) = s_1^m$, $g_1^{-1}(q) = g^{-1}(q) = s_2^m + Q^{m-1}$, $\gamma_q^m(g_1) = \gamma_q^m(g) = \alpha_2\sigma_2^m$. Since s_1^m is a sphere and $\gamma_p^m(g_1) \subset s_1^m$, therefore $\gamma_p^m(g_1) = \alpha\sigma_1^m$. It follows that $c(g_1) = v[\gamma_p^m(g_1), \gamma_q^m(g_1)] = v[\alpha\sigma_1^m, \alpha_2\sigma_2^m] = \alpha\alpha_2$. On the other hand $c(g_1) = c(g) = \alpha_1\alpha_2$ and therefore $\alpha = \alpha_1$. We have proved therefore that $\gamma_p^m(g_1) = \gamma_p^m(g_2) = \alpha_1\sigma_1^m$. Repeating the same argument with g replaced by g_1 and p replaced by q we obtain a map g_2 homotopic to g_1 and such that

$$g_2^{-1}(p) = s_1^m, \qquad \gamma_p^m(g_2) = \alpha_1\sigma_1^m,$$

$$g_2^{-1}(q) = s_2^m, \qquad \gamma_q^m(g_2) = \alpha_2\sigma_2^m.$$

Sufficiency follows from Th. III.

V. Mappings of S^{n+1} into S^{m+1}

1. Lemma 5. *Given a regular mapping $f(M^{n+1}) \subset S^{m+1}$ and a decomposition $M^{n+1} = X_1 + X_2$ into two closed sets such that $f(X_1) \subset S^{m+1} - q$, $f(X_2) \subset S^{m+1} - p$, there is a mapping g homotopic to f such that*

$$g^{-1}(p) = f^{-1}(p), \qquad \gamma_p^{n-m}(g) = \gamma_p^{n-m}(f),$$

$$g^{-1}(q) = f^{-1}(q), \qquad \gamma_q^{n-m}(g) = \gamma_q^{n-m}(f),$$

$$g(X_1) \subset E_+^{m+1}, \qquad g(X_2) \subset E_-^{m+1}.$$

Proof. Let $\eta < 1$ be a number such that $x \in f(X_1)$ implies $l(x) > -\eta$ and $x \in f(X_2)$ implies $l(x) < \eta$.

Let $k(x)$ be a continuous mapping of the interval $-1, 1$ into itself such that

$$k(x) = x \quad \text{if} \quad x \geqq \tfrac{1}{2}(\eta + 1),$$

$$k(x) = 0 \quad \text{if} \quad x \leqq \eta,$$

$$k(x) \leqq x \quad \text{for all} \quad 1 \leqq x \leqq 1.$$

We define a map $s(S^{m+1}) \subset S^{m+1}$ by taking $s(p) = p$, $s(q) = q$, if $x = \{r(x), l(x)\}$ then $s(x) = \{r(x), k[l(x)]\}$. The map $g(x) = s[f(x)]$ satisfies the conditions of Lemma 5.

2. *Proof of* Th. Ib *and* IIb. By Th. Ia and IIa it is enough to prove that given a map $f(S^{n+1}) \subset S^{m+1}$ such that $f^{-1}(p) = p$, $f^{-1}(q) = q$ there is a map $\varphi(S^n) \subset S^m$ such that F^φ is homotopic to f.

Since $f(E_+^{n+1}) \subset S^{m+1} - q$ and $f(E_-^{n+1}) \subset S^{m+1} - p$ there is by Lemma 5 a map g homotopic to f and such that $g(E_+^{n+1}) \subset E_+^{n+1}$, $g(E_-^{n+1}) \subset E_-^{n+1}$. It follows that $g(S^n) \subset S^m$. Let $\varphi(S^n) \subset S^m$ be the map of S^n thus defined. It is easy to see that for each $x \in S^{n+1}$ we have $| g(x) - F^\varphi(x) | < 2$ and therefore that g and F^φ are homotopic.

3. *Proof of* Th. IIIb. Since $0 \times S^m$ and $S^m \times 0$ form an elementary couple in $E^{m+1} \times S^m + S^m \times E^{m+1}$ therefore by Theorem IIIa it is enough to prove that each map $f(E^{m+1} \times S^m + S^m \times E^{m+1}) \subset S^{m+1}$ such that

$$f^{-1}(p) = 0 \times S^m, \qquad \gamma_p^m(f) = \alpha_1(0 \times \sigma^m),$$

$$f^{-1}(q) = S^m \times 0, \qquad \gamma_q^m(f) = \alpha_2(\sigma^m \times 0).$$

(where σ^m is a basis cycle for S^m) is homotopic to a map H^φ where $\varphi(S^m \times S^m) \subset S^m$ is a map of the type (α_1, α_2).

Since $f(E^{m+1} \times S^m) \subset S^{m+1} - q$ and $f(S^m \times E^{m+1}) \subset S^{m+1} - p$ therefore by Lemma 5 there is a map g homotopic to f such that

$$g^{-1}(p) = 0 \times S^m, \qquad \gamma_p^m(g) = \alpha_1(0 \times \sigma^m),$$

$$g^{-1}(q) = S^m \times 0, \qquad \gamma_q^m(g) = \alpha_2(\sigma^m \times 0),$$

$$g(E^{m+1} \times S^m) \subset E_+^{m+1}, \qquad g(S^m \times E^{m+1}) \subset E_-^{m+1}.$$

It follows that $g(S^m \times S^m) \subset S^m$. Let $\varphi(S^m \times S^m) \subset S^m$ be the map of $S^m \times S^m$ thus defined. We then have $| g(x) - H^\varphi(x) | < 2$ for each $x \in E^{m+1} \times S^m + S^m \times E^{m+1}$ and therefore g and H^φ are homotopic.

For each $t \in S^m$ the element $E^{m+1} \times t$ has an intersection coefficient 1 with $0 \times \sigma^m$ and therefore an intersection coefficient α_1 with $\gamma_p^m(g) = \alpha_1(0 \times \sigma^m)$. It follows by (1) that $\mathfrak{g}[g, E^{m+1} \times t, p] = \alpha_1$ and therefore $\mathfrak{g}[\varphi, S^m \times t] = \alpha_1$. Similarly we obtain $\mathfrak{g}[\varphi, t \times S^m] = \alpha_2$. Therefore φ is of the type (α_1, α_2).

UNIVERSITY OF MICHIGAN.

obtain therefore a map g_1 homotopic to g such that $g_1^{-1}(p) = s_1^m$, $g_1^{-1}(q) = g^{-1}(q) = s_2^m + Q^{m-1}$, $\gamma_q^m(g_1) = \gamma_q^m(g) = \alpha_2\sigma_2^m$. Since s_1^m is a sphere and $\gamma_p^m(g_1) \subset s_1^m$, therefore $\gamma_p^m(g_1) = \alpha\sigma_1^m$. It follows that $c(g_1) = v[\gamma_p^m(g_1), \gamma_q^m(g_1)] = v[\alpha\sigma_1^m, \alpha_2\sigma_2^m] = \alpha\alpha_2$. On the other hand $c(g_1) = c(g) = \alpha_1\alpha_2$ and therefore $\alpha = \alpha_1$. We have proved therefore that $\gamma_p^m(g_1) = \gamma_p^m(g_2) = \alpha_1\sigma_1^m$. Repeating the same argument with g replaced by g_1 and p replaced by q we obtain a map g_2 homotopic to g_1 and such that

$$g_2^{-1}(p) = s_1^m, \qquad \gamma_p^m(g_2) = \alpha_1\sigma_1^m,$$

$$g_2^{-1}(q) = s_2^m, \qquad \gamma_q^m(g_2) = \alpha_2\sigma_2^m.$$

Sufficiency follows from Th. III.

V. Mappings of S^{n+1} into S^{m+1}

1. LEMMA 5. *Given a regular mapping $f(M^{n+1}) \subset S^{m+1}$ and a decomposition $M^{n+1} = X_1 + X_2$ into two closed sets such that $f(X_1) \subset S^{m+1} - q$, $f(X_2) \subset S^{m+1} - p$, there is a mapping g homotopic to f such that*

$$g^{-1}(p) = f^{-1}(p), \qquad \gamma_p^{n-m}(g) = \gamma_p^{n-m}(f),$$

$$g^{-1}(q) = f^{-1}(q), \qquad \gamma_q^{n-m}(g) = \gamma_q^{n-m}(f),$$

$$g(X_1) \subset E_+^{m+1}, \qquad g(X_2) \subset E_-^{m+1}.$$

PROOF. Let $\eta < 1$ be a number such that $x \in f(X_1)$ implies $l(x) > -\eta$ and $x \in f(X_2)$ implies $l(x) < \eta$.

Let $k(x)$ be a continuous mapping of the interval $-1, 1$ into itself such that

$$k(x) = x \quad \text{if} \quad x \geqq \tfrac{1}{2}(\eta + 1),$$

$$k(x) = 0 \quad \text{if} \quad x \leqq \eta,$$

$$k(x) \leqq x \quad \text{for all} \quad 1 \leqq x \leqq 1.$$

We define a map $s(S^{m+1}) \subset S^{m+1}$ by taking $s(p) = p$, $s(q) = q$, if $x = \{r(x), l(x)\}$ then $s(x) = \{r(x), k[l(x)]\}$. The map $g(x) = s[f(x)]$ satisfies the conditions of Lemma 5.

2. *Proof of* Th. Ib *and* IIb. By Th. Ia and IIa it is enough to prove that given a map $f(S^{n+1}) \subset S^{m+1}$ such that $f^{-1}(p) = p$, $f^{-1}(q) = q$ there is a map $\varphi(S^n) \subset S^m$ such that F^φ is homotopic to f.

Since $f(E_+^{n+1}) \subset S^{m+1} - q$ and $f(E_-^{n+1}) \subset S^{m+1} - p$ there is by Lemma 5 a map g homotopic to f and such that $g(E_+^{n+1}) \subset E_+^{n+1}$, $g(E_-^{n+1}) \subset E_-^{n+1}$. It follows that $g(S^n) \subset S^m$. Let $\varphi(S^n) \subset S^m$ be the map of S^n thus defined. It is easy to see that for each $x \in S^{n+1}$ we have $| g(x) - F^\varphi(x) | < 2$ and therefore that g and F^φ are homotopic.

3. *Proof of* Th. IIIb. Since $0 \times S^m$ and $S^m \times 0$ form an elementary couple in $E^{m+1} \times S^m + S^m \times E^{m+1}$ therefore by Theorem IIIa it is enough to prove that each map $f(E^{m+1} \times S^m + S^m \times E^{m+1}) \subset S^{m+1}$ such that

$$f^{-1}(p) = 0 \times S^m, \qquad \gamma_p^m(f) = \alpha_1(0 \times \sigma^m),$$

$$f^{-1}(q) = S^m \times 0, \qquad \gamma_q^m(f) = \alpha_2(\sigma^m \times 0).$$

(where σ^m is a basis cycle for S^m) is homotopic to a map H^φ where $\varphi(S^m \times S^m) \subset S^m$ is a map of the type (α_1, α_2).

Since $f(E^{m+1} \times S^m) \subset S^{m+1} - q$ and $f(S^m \times E^{m+1}) \subset S^{m+1} - p$ therefore by Lemma 5 there is a map g homotopic to f such that

$$g^{-1}(p) = 0 \times S^m, \qquad \gamma_p^m(g) = \alpha_1(0 \times \sigma^m),$$

$$g^{-1}(q) = S^m \times 0, \qquad \gamma_q^m(g) = \alpha_2(\sigma^m \times 0),$$

$$g(E^{m+1} \times S^m) \subset E_+^{m+1}, \qquad g(S^m \times E^{m+1}) \subset E_-^{m+1}.$$

It follows that $g(S^m \times S^m) \subset S^m$. Let $\varphi(S^m \times S^m) \subset S^m$ be the map of $S^m \times S^m$ thus defined. We then have $| g(x) - H^\varphi(x) | < 2$ for each $x \in E^{m+1} \times S^m + S^m \times E^{m+1}$ and therefore g and H^φ are homotopic.

For each $t \in S^m$ the element $E^{m+1} \times t$ has an intersection coefficient 1 with $0 \times \sigma^m$ and therefore an intersection coefficient α_1 with $\gamma_p^m(g) = \alpha_1(0 \times \sigma^m)$. It follows by (1) that $\mathfrak{g}[g, E^{m+1} \times t, p] = \alpha_1$ and therefore $\mathfrak{g}[\varphi, S^m \times t] = \alpha_1$. Similarly we obtain $\mathfrak{g}[\varphi, t \times S^m] = \alpha_2$. Therefore φ is of the type (α_1, α_2).

UNIVERSITY OF MICHIGAN.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

ON p-ADIC FIELDS AND RATIONAL DIVISION ALGEBRAS[1]

BY A. A. ALBERT

(Received January 20, 1940)

1. Introduction

The known results[2] on the existence of abelian extensions, with prescribed properties, of an algebraic number field have been obtained as a part of the class field theory. This latter theory is certainly of great importance but is undeniably of tremendous complexity. It would thus be very desirable, if possible, to obtain a simpler and more algebraic derivation of the theorems on abelian extensions mentioned above.

The corresponding theorems on abelian extensions of p-adic number fields have been obtained[3] as analogues, in the local class field theory, of those on extensions of algebraic number fields. This treatment is again undesirable, and moreover, is at least partly unnecessary. The class field theoretic treatment also hides the really simple results which it is possible to obtain on the structure of arbitrary fields whose ramification order is prime to p, as well as the fact that abelian extensions of a prescribed degree over a p-adic field $\mathfrak{F}$ may not even exist.

We shall use a very simple lemma[4] of what may be called an arithmetical nature on p-adic fields and shall give a consequent new algebraic discussion of fields $\mathfrak{K}$ of finite degree and ramification order e prime to p over a p-adic field $\mathfrak{F}$. We shall construct all such fields, show that the number of such fields inequivalent over $\mathfrak{F}$ is an explicitly determined divisor of e, and shall determine necessary and sufficient conditions that $\mathfrak{K}$ be normal over $\mathfrak{F}$. We shall determine the automorphism group of $\mathfrak{K}$ over $\mathfrak{F}$, and prove the rather amazing result that the existence of any abelian field of the given ramification order e over $\mathfrak{F}$ implies that all fields with the same e are abelian over $\mathfrak{F}$. Finally we shall determine explicitly which abelian fields are cyclic, how many cyclic fields exist which are inequivalent over $\mathfrak{F}$, the direct factorization into two cyclic fields of all non-cyclic abelian fields.

An extension of our treatment to abelian fields of ramification order not prime to p over a p-adic field might be obtainable but would probably be so complicated

---

[1] Presented to the Society December 28, 1939.

[2] See [7] of the bibliography at the end of this article.

[3] In [8], [28]. See also the theory of class fields as given in [10], [5], [9], and of local class fields in [6], [21], [24], [25].

[4] A more general form of this lemma is due to G. E. Wahlin and was first proved by him in [29]. We use this more general form later in considering fields of arbitrary ramification order, but believe it desirable to make the theory of p-adic fields of ramification order prime to p depend upon as little p-adic theory as possible. Hence we indicate the very simple proof of this more elementary lemma.

that it might be just another form of that given by the local class field theory. However some results are obtainable simply. We shall consider fields $\mathfrak{K}$ of residue-class degree f and ramification order e over a p-adic field $\mathfrak{F}$ and write $e = p^{\nu}e_0$ with e_0 prime to p. We shall then show that $\mathfrak{K}$ is completely ramified of degree p^{ν} over a subfield of ramification order e_0 and residue-class degree f over $\mathfrak{F}$. We shall use this result to give new simple proofs of the theorems[5] stating that all normal extensions of a p-adic field $\mathfrak{F}$ are solvable, and that there are only a finite number of extensions of a given degree inequivalent over $\mathfrak{F}$.

It is known[6] that if $\mathfrak{D}$ is a generalized quaternion division algebra over a quadratic extension $\mathfrak{K}$ of the rational number field $\mathfrak{R}$ then $\mathfrak{D}$ has a maximal subfield $\mathfrak{K}(x)$, $x^2 = a$ in $\mathfrak{R}$. To obtain a generalization of this result we may consider a normal division algebra $\mathfrak{D}$ of degree m over its centrum $\mathfrak{K}$ and let $\mathfrak{K}$ have degree r over an algebraic number field $\mathfrak{F}$. Then we ask whether or not there exists a maximal subfield $\mathfrak{K}(x)$ of $\mathfrak{D}$ such that the minimum function of x has coefficients in $\mathfrak{F}$ and the field $\mathfrak{F}(x)$ is cyclic over $\mathfrak{F}$. This question is equivalent to that about the existence of a cyclic field $\mathfrak{Z}$ of degree m over $\mathfrak{F}$ such that the composite of $\mathfrak{Z}$ and $\mathfrak{K}$ splits $\mathfrak{D}$. We shall show that $\mathfrak{Z}$ may, in general, not exist. We shall, indeed, find necessary and sufficient conditions that $\mathfrak{Z}$ exist. In particular $\mathfrak{Z}$ always will exist if m is prime to r.

2. Notations and elementary results[7]

Let p be a rational prime integer which will be fixed until we say otherwise. We speak of the derived field $\mathfrak{R}_p$ of the field $\mathfrak{R}$ of all rational numbers as the field of rational p-adic numbers. If $\mathfrak{F}$ is any algebraic extension of $\mathfrak{R}_p$ we call the quantities of $\mathfrak{F}$ *p-adic algebraic numbers*.

We shall consider a field $\mathfrak{F}$ of finite degree over $\mathfrak{R}_p$. The set $J_{\mathfrak{F}}$ of all integers of $\mathfrak{F}$ has a unique prime ideal $\mathfrak{p}_{\mathfrak{F}}$ and we let $q_{\mathfrak{F}}$ be the number of elements in the residue class field $H_{\mathfrak{F}} = J_{\mathfrak{F}} - \mathfrak{p}_{\mathfrak{F}}$. This field $H_{\mathfrak{F}}$ is then a finite (Galois) field $GF(q_{\mathfrak{F}})$, where $q_{\mathfrak{F}}$ is a power of p.

We shall make our discussion relative to a fixed prime integer $P_{\mathfrak{F}}$ of $\mathfrak{F}$. Then $\mathfrak{p}_{\mathfrak{F}} = (P_{\mathfrak{F}})$, every prime integer of $\mathfrak{F}$ has the form $uP_{\mathfrak{F}}$ where u is a unit of $J_{\mathfrak{F}}$. We shall speak of the units of the integral domain $J_{\mathfrak{F}}$ simply as *units* of $\mathfrak{F}$, the prime integers of $\mathfrak{F}$ as *primes* or *prime quantities* of $\mathfrak{F}$.

Every p-adic field $\mathfrak{F}$ contains a primitive $(q_{\mathfrak{F}} - 1)^{\text{th}}$ root of unity which we

---

[5] The first of these theorems is due to Hensel and is stated in an equivalent form and proved in **[16]** by the use of the method developed by Hilbert in his theory of ramification fields. The second theorem is stated by Krasner in **[18]** with methods of computing the number of fields in **[19]**, **[20]**. He indicates without details that the theorem itself follows from the results of **[17]**.

[6] This is a result in **[1]**. It is generalized to algebras over an algebraic number field in **[2]**.

[7] We use the exposition of p-adic theory of **[3]**, **[4]** as a basis for our present theory. Note that the very interesting results of Hensel in **[11]**-**[16]** could be simplified greatly if based upon the more modern exposition of these two texts, due essentially to Hensel, H. Hasse, and O. Ostrowski.

shall fix throughout our discussion and designate by $\zeta_{\mathfrak{F}}$. Every quantity a of $\mathfrak{F}$ is expressible in the form

$$a = P_{\mathfrak{F}}^{\lambda}\zeta_{\mathfrak{F}}^{\mu}u, \tag{1}$$

where λ and μ are rational integers and u is a unit of $\mathfrak{F}$ such that

$$u \equiv 1 \pmod{\mathfrak{p}_{\mathfrak{F}}}. \tag{2}$$

The expression of a in the form (1) is unique, that is, if also $a = P_{\mathfrak{F}}^{\alpha}\zeta_{\mathfrak{F}}^{\beta}u_0$ for $u_0 \equiv 1 \pmod{\mathfrak{p}_{\mathfrak{F}}}$ then $\lambda = \alpha$, $\mu = \beta$, $u = u_0$. The integer λ is called the *order* of a.

The integers of $\mathfrak{F}$ are those quantities (1) with $\lambda \geqq 0$, and the units of $\mathfrak{F}$ are the quantities with $\lambda = 0$. We call units of $\mathfrak{F}$ satisfying (2) *principal* units of $\mathfrak{F}$. The set of all such units forms a subgroup of the multiplicative group of all units of $\mathfrak{F}$.

We now let $\mathfrak{K}$ be a field of degree n over a p-adic field $\mathfrak{F}$. The meaning of

$$J_{\mathfrak{K}}, \mathfrak{p}_{\mathfrak{K}}, q_{\mathfrak{K}}, H_{\mathfrak{K}} = J_{\mathfrak{K}} - \mathfrak{p}_{\mathfrak{K}} \tag{3}$$

of $\mathfrak{K}$ is clear, and we have

$$n = ef, \qquad \mathfrak{p}_{\mathfrak{K}} = \mathfrak{p}_{\mathfrak{F}}^{e}, \qquad q_{\mathfrak{K}} = q_{\mathfrak{F}}^{f}, \tag{4}$$

where e is the ramification order of $\mathfrak{K}$ over $\mathfrak{F}$, f is the residue-class degree of $\mathfrak{K}$ over $\mathfrak{F}$. The field $\mathfrak{K}$ contains a primitive $(q_{\mathfrak{K}} - 1)^{\text{th}}$ root of unity $\zeta_{\mathfrak{K}}$, and we shall write

$$\tau = q_{\mathfrak{K}} - 1 = (q_{\mathfrak{F}} - 1)\sigma, \qquad \sigma = 1 + q_{\mathfrak{F}} + \cdots + q_{\mathfrak{F}}^{f-1}, \tag{5}$$

and shall assume throughout our discussion that

$$\zeta_{\mathfrak{F}} = \zeta_{\mathfrak{K}}^{\sigma}. \tag{6}$$

If now $P_{\mathfrak{K}}$ is any prime quantity of $\mathfrak{K}$ the expression of any A of $\mathfrak{K}$ in the form

$$A = P_{\mathfrak{K}}^{\lambda}\zeta_{\mathfrak{K}}^{\mu}U \tag{7}$$

for rational integral λ and μ, and U a principal unit of $\mathfrak{K}$, is unique.

We note that every principal unit $U \neq 1$ of $\mathfrak{K}$ has the form $U = 1 + P_{\mathfrak{K}}^{\lambda_0}U_0$ for U_0 a unit of $\mathfrak{K}$, $\lambda_0 > 0$. We then prove

LEMMA 1. *Let m be a rational integer prime to p. Then 1 is the only principal unit of $\mathfrak{K}$ which is an m^{th} root of unity.*

For if $U \neq 1$ then $U = 1 + P_{\mathfrak{K}}^{\lambda_0}U_0$, mU_0 is a unit U_1 of $\mathfrak{K}$, $U^m = 1 + mU_0P_{\mathfrak{K}}^{\lambda_0} + P_{\mathfrak{K}}^{2\lambda_0}B$ where B is an integer of $\mathfrak{K}$. Then $U^m - 1 = P_{\mathfrak{K}}^{\lambda_0}(U_1 + P^{\lambda_0}B) \neq 0$.

As a simple consequence of the lemma just proved we have

LEMMA 2. *Let m be prime to p and A in $\mathfrak{K}$ be such that $A^m = P_{\mathfrak{K}}^{\lambda}\zeta_{\mathfrak{K}}^{\mu}$. Then m divides $\lambda = \lambda_0 m$ and $\mu = \mu_0 m$, and $A = P_{\mathfrak{K}}^{\lambda_0}\zeta_{\mathfrak{K}}^{\mu_0}$.*

For $A = P_{\mathfrak{K}}^{\lambda_0}\zeta_{\mathfrak{K}}^{\mu_0}U$ for a principal unit U of $\mathfrak{K}$. Then $A^m = P_{\mathfrak{K}}^{m\lambda_0}\zeta_{\mathfrak{K}}^{m\mu_0}U^m = P_{\mathfrak{K}}^{\lambda}\zeta_{\mathfrak{K}}^{\mu}$. Since U^m is a principal unit of $\mathfrak{K}$ we have $m\lambda_0 = \lambda$, $m\mu_0 = \mu$, $U^m = 1$. By Lemma 1 we have $U = 1$.

The case $\lambda = 0$ of the result above gives

LEMMA 3. *Let m be prime to p and U be in $\mathfrak{K}$ and such that U^m is a power of $\zeta_{\mathfrak{K}}$. Then U is a power of $\zeta_{\mathfrak{K}}$.*

We also have

LEMMA 4. *Let m be prime to p. Then K contains a primitive m^{th} root of unity ζ if and only if m divides $\tau = q_{\mathfrak{K}} - 1$, ζ is a power of $\zeta_{\mathfrak{K}}$.*

For if $\zeta = \zeta_{\mathfrak{K}}^r$ and $\tau = mr$ the quantity ζ is a primitive m^{th} root of unity in $\mathfrak{K}$. Conversely if ζ in $\mathfrak{K}$ is a primitive m^{th} root of unity we have $\zeta^m = 1 = \zeta_{\mathfrak{K}}^\tau$ and Lemma 3 implies that $\zeta = \zeta_{\mathfrak{K}}^r$. The cyclic multiplicative group generated by $\zeta_{\mathfrak{K}}$ has order τ and has the cyclic subgroup of order m generated by ζ, m divides τ.

Our final preliminary lemma will be found to be fundamental in our later discussion.[8]

LEMMA 5. *If m is prime to p and U is a principal unit of $\mathfrak{K}$ there exists a principal unit V in $\mathfrak{K}$ such that $V^m = U$.*

For let $f(x) = x^m - U$. Since $U \equiv 1 \pmod{\mathfrak{p}_{\mathfrak{K}}}$ we have $f(x) \equiv x^m - 1 \equiv (x-1)g_0(x) \pmod{\mathfrak{p}_{\mathfrak{K}}}$. Here $g_0(x)$ is a polynomial of degree $m - 1$ of $H_{\mathfrak{K}}[x]$ and is prime to $x - 1$. For $x^m - 1$ is separable when m is prime to the characteristic p of $H_{\mathfrak{K}}$. By *Hensel's Lemma* there exists a quantity V in $J_{\mathfrak{K}}$ such that $f(x) = (x - V)g(x)$, $V \equiv 1 \pmod{\mathfrak{p}_{\mathfrak{K}}}$. But then $f(V) = 0$, V is a principal unit of $\mathfrak{K}$ such that $V^m = U$.

The properties derived above are not properties of $\mathfrak{K}$ relative to $\mathfrak{F}$ but are what we may call *absolute* properties of $\mathfrak{K}$. They then hold for any subfield $\mathfrak{F}$ over R_p of $\mathfrak{K}$. However, we shall also require some elementary properties of $\mathfrak{K}$ relative to $\mathfrak{F}$.

The field $\mathfrak{W} = \mathfrak{F}(\zeta_{\mathfrak{K}})$ is unramified of degree f over $\mathfrak{F}$, that is, $P_{\mathfrak{F}}$ is a prime quantity of $\mathfrak{W}$. Also $\zeta_{\mathfrak{K}} = \zeta_{\mathfrak{W}}$. If $\mathfrak{W}_0$ is any unramified subfield over $\mathfrak{F}$ of $\mathfrak{K}$ its residue class degree over $\mathfrak{F}$ is a divisor f_0 of f, the degree f_0 of $\mathfrak{W}_0$ over $\mathfrak{F}$ divides the degree of $\mathfrak{W}$, $\mathfrak{W}_0$ is equivalent over $\mathfrak{F}$ to a subfield $\mathfrak{W}_1$ of $\mathfrak{W}$. But $\mathfrak{W}_1$ is normal over $\mathfrak{F}$, $\mathfrak{W}_1$ contains all subfields of $\mathfrak{K}$ which are equivalent over $\mathfrak{F}$ to itself, $\mathfrak{W}_0 = \mathfrak{W}_1 \leqq \mathfrak{W}$. We have shown that $\mathfrak{W}$ is the *maximal unramified subfield* over $\mathfrak{F}$ of $\mathfrak{K}$ in the sense that *it contains all subfields of $\mathfrak{K}$ unramified over $\mathfrak{F}$.*

A well known theorem[9] states that $\mathfrak{K} = \mathfrak{W}(P)$ for any prime quantity P of $\mathfrak{K}$ and that $\mathfrak{K}$ is completely ramified of degree e over $\mathfrak{F}$. Now $(P)^e = \mathfrak{p}_{\mathfrak{F}} = (P_{\mathfrak{F}})$, $P^e = UP_{\mathfrak{F}}$ for a unit U of $\mathfrak{K}$. Since $\mathfrak{K} = \mathfrak{W}(P)$ we have $U = a_0 + a_1P + \cdots + a_{e-1}P^{e-1}$ for a_i in $\mathfrak{W}$. Then P is a root of

$$f(x) = x^e - P_{\mathfrak{F}}a_{e-1}x^{e-1} - \cdots - P_{\mathfrak{F}}a_1x - P_{\mathfrak{F}}a_0 = 0. \tag{8}$$

We conclude that $f(x)$ must be the minimum function over $\mathfrak{W}$ of P and must be irreducible in $\mathfrak{W}$. But P is an integer of $\mathfrak{K}$ and the coefficients of $f(x)$ are integers $b_i = P_{\mathfrak{F}}a_i$ of $\mathfrak{W}$. The residue-class degree of $\mathfrak{K}$ over $\mathfrak{W}$ is unity, $N_{\mathfrak{K}|\mathfrak{W}}(P) = \pm a_0P_{\mathfrak{F}}$ has order unity, a_0 is a unit of $\mathfrak{W}$. The ideal $(P)^e = \mathfrak{p}_{\mathfrak{W}}$

[8] This is the lemma referred to in our introduction.

[9] See Chapter IX of [4].

divides $P^e - P_{\mathfrak{F}}a_0 = b_1P + P^2(b_2 + \cdots + b_{e-1}P^{e-3})$, $(P)^2$ divides b_1P, (P) divides b_1 and b_1 cannot be a unit of $\mathfrak{W}$. Hence $b_1 = P_{\mathfrak{F}}a_1$ for a_1 an integer of $\mathfrak{W}$. Continuing in this fashion we ultimately prove that all the a_i are integers of $\mathfrak{W}$. This is the first part of[10]

THEOREM 1. *Let $\mathfrak{W}$ be the maximal unramified subfield over $\mathfrak{F}$ of $\mathfrak{K}$ over $\mathfrak{F}$, P be any prime quantity of $\mathfrak{K}$, e be the ramification order over $\mathfrak{F}$ of $\mathfrak{K}$, $P_{\mathfrak{F}}$ be a prime quantity of $\mathfrak{F}$. Then the minimum function over $\mathfrak{W}$ of P is*

$$f(x) = x^e + P_{\mathfrak{F}}(c_0 + c_1x + \cdots + c_{e-1}x^{e-1}), \tag{9}$$

for integers c_i of $\mathfrak{W}$, c_0 a unit of $\mathfrak{W}$. Conversely if c_0 is a unit and the remaining c_i are integers of a field $\mathfrak{W}$ unramified over $\mathfrak{F}$ any stem field $\mathfrak{K} = \mathfrak{W}(\xi)$ of (9) has ramification order e over $\mathfrak{F}$ and $\mathfrak{W}$ as maximal unramified subfield, ξ is a prime quantity of $\mathfrak{K}$, $U = c_0 + c_1\xi + \cdots + c_{e-1}\xi^{e-1}$ is a unit of $\mathfrak{K}$.

For the proof of the second part of our theorem we note that any root ξ of $f(x) = 0$ is an integer of $\mathfrak{K} = \mathfrak{F}(\xi)$. Then so is U, $\xi^e = -P_{\mathfrak{F}}U$ must be divisible by $\mathfrak{p}_{\mathfrak{K}}$. But c_0 is a unit of $\mathfrak{W}$ and of $\mathfrak{K}$ and is not divisible by $\mathfrak{p}_{\mathfrak{K}}$, $U = c_0 + \xi A$ is not divisible by $\mathfrak{p}_{\mathfrak{K}}$ and is a unit of $\mathfrak{K}$, the ideal $(\xi)^e = (-UP_{\mathfrak{F}}) = \mathfrak{p}_{\mathfrak{W}}$. Hence the ramification order of $\mathfrak{K}$ over $\mathfrak{W}$ is at least e, the degree of $\mathfrak{K}$ over $\mathfrak{W}$ is at most e, $\mathfrak{K}$ has degree and ramification order e over $\mathfrak{W}$. Evidently then $f(x)$ is irreducible in $\mathfrak{W}$, $(\xi) = \mathfrak{p}_{\mathfrak{K}}$, ξ is a prime quantity of $\mathfrak{F}$. Since $\mathfrak{K}$ is completely ramified over $\mathfrak{W}$ the field $\mathfrak{W}$ is the maximal unramified subfield over $\mathfrak{F}$ of $\mathfrak{K}$.

3. Fields with e prime to p

The structure of algebraic extensions of ramification order e prime to p over a p-adic field $\mathfrak{F}$ is remarkably simple. Our results are fundamentally due to two simple theorems, the first of which is

THEOREM 2. *Let $\mathfrak{F}$ be a p-adic field, e and f be rational integers,*

$$\tau = q_{\mathfrak{F}}^f - 1, \qquad (e, p) = 1, \qquad \mathfrak{W} = \mathfrak{F}(\zeta_{\mathfrak{W}}) \tag{10}$$

for a primitive τ^{th} root of unity $\zeta_{\mathfrak{W}}$. Then every field $\mathfrak{K}$ of degree ef and ramification order e over $\mathfrak{F}$ is equivalent to

$$\mathfrak{K} = \mathfrak{W}(P), \qquad P^e = \zeta_{\mathfrak{W}}^i P_{\mathfrak{F}}, \tag{11}$$

for a rational integer i which we shall call a **type number** *of $\mathfrak{K}$. Conversely the fields (11) have ramification order e and degree ef over $\mathfrak{F}$.*

For by Theorem 1 we may write $\mathfrak{K} = \mathfrak{W}(P)$, $P^e = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i U$ for U a principal unit of $\mathfrak{K}$. By Lemma 5 there exists a principal unit V in $\mathfrak{K}$ such that $V^e = U$. Write $P_0 = V^{-1}P$ so that P_0 is a prime quantity of $\mathfrak{K}$, $\mathfrak{K} = \mathfrak{W}(P_0)$, $P_0^e = U^{-1}P^e = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i$ as desired. The converse is an instance of Theorem 1. Note our use of (a, b) for the greatest common divisor of two rational integers a and b.

We next have

THEOREM 3. *Let $\mathfrak{K}$ and $\mathfrak{L}$ be two fields of the same degree and ramification order*

[10] This is a type of Eisenstein theorem. See in this connection p. 74 of [**11**].

e over $\mathfrak{F}$ $(e, p) = 1$, and let i and j be a pair of respective type numbers. Then $\mathfrak{K}$ and $\mathfrak{L}$ are equivalent over $\mathfrak{F}$ if and only if i and j are congruent modulo

$$d = (\tau, e). \tag{12}$$

Thus the fields

$$\mathfrak{K}_i = \mathfrak{W}(P_i), \qquad P_i^e = \zeta_{\mathfrak{W}}^i P_{\mathfrak{F}} \quad (i = 0, 1, \cdots, d-1), \tag{13}$$

give a set of d inequivalent fields of the same degree ef and ramification order e prime to p over $\mathfrak{F}$ to one of which every such field[11] *is equivalent.*

For if $\mathfrak{K} = \mathfrak{W}(P)$, $P^e = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^{\lambda}$ we write $\lambda = dk + i$ $(0 \leqq i < d)$, $d = ee_1 + \tau\tau_1$, for rational integers e_1 and τ_1, $\zeta_{\mathfrak{W}}^d = (\zeta_{\mathfrak{W}}^{e_1})^e$, $\zeta_{\mathfrak{W}}^{\lambda} = (\zeta_{\mathfrak{W}}^{e_1k})^e\zeta_{\mathfrak{W}}^i$. The quantity $P_i = \zeta_{\mathfrak{W}}^{-e_1k} P$ is a prime quantity of $\mathfrak{K}$, $\mathfrak{K} = \mathfrak{W}(P_i)$ as in (13). Conversely every field (11) has degree ef and ramification order e over $\mathfrak{F}$, this is true in particular of the fields (13), the fields (11) are equivalent to fields (13). If $\mathfrak{K}$ has a type number i and $\mathfrak{L}$ has a type number j then $\mathfrak{K}$ and $\mathfrak{L}$ are equivalent over $\mathfrak{F}$ if and only if $\mathfrak{K}$ contains a prime quantity P_0 such that $P_0^e = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^j$. Now $P_0 = UP$ for a unit U of $\mathfrak{K}$, $P_0^e = U^eP^e = U^eP_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^j$. Then $U^e = \zeta_{\mathfrak{W}}^{j-i}$. By Lemma 3 $U = \zeta_{\mathfrak{W}}^k$, $\zeta_{\mathfrak{W}}^{ke} = \zeta_{\mathfrak{W}}^{j-i}$ which is true if and only if

$$j - i \equiv ke \pmod{\tau}. \tag{14}$$

But this congruence has a solution k if and only if d divides $j - i$.

We note the corollary of our proof.

LEMMA 6. *Let S be any automorphism over F of a field K of* (11). *Then $P^S = \zeta_{\mathfrak{W}}^k P$.*

For $\zeta_{\mathfrak{W}}^S$ is a primitive τ^{th} root of unity and is a power of $\zeta_{\mathfrak{W}}$, $(P^S)^e = \zeta_{\mathfrak{W}}^S P_{\mathfrak{F}}^S = \zeta_{\mathfrak{W}}^S P_{\mathfrak{F}}$. Our result follows from the proof above.

We may however obtain the more explicit result of

THEOREM 4. *The automorphisms S over F of a field K of* (11) *carry $\zeta_{\mathfrak{W}}$ into $\zeta_{\mathfrak{W}}^{s_\mu}$ where*

$$s_\mu = q_{\mathfrak{F}}^{\mu}, \tag{15}$$

for a rational integer μ. If then S is any automorphism over $\mathfrak{F}$ of $\mathfrak{K}$ we have

$$P^S = \zeta_{\mathfrak{W}}^k P, \qquad \zeta_{\mathfrak{W}}^S = \zeta_{\mathfrak{W}}^{s_\mu} \tag{16}$$

for some integer μ and an integer k such that

$$ke \equiv i(q_{\mathfrak{F}}^{\mu} - 1) \pmod{\tau}. \tag{17}$$

Conversely the correspondence in $\mathfrak{K}$ over $\mathfrak{F}$ induced by (16) *is an automorphism S*

---

[11] It is natural at this point to discuss also the subfields of a given $\mathfrak{K}$, and the number of inequivalent subfields of a given ramification order and residue-class degree. The author has studied and solved this problem but has not included the results in the present paper because to do so would delay unnecessarily our real goal, the study of abelian fields. The methods of obtaining these additional results is that of the proof of the present theorem.

over F of $\mathfrak{K}$ for every solution k of (17), *S exists if and only if $d = (e, \tau)$ divides* $i(q_{\mathfrak{F}} - 1)$.

For Lemma 6 states that $(P^S)^e = \zeta_{\mathfrak{W}}^{ke} P^e = \zeta_{\mathfrak{W}}^{ke+i} P_{\mathfrak{F}}$ while clearly we must have $(P^S)^e = P_{\mathfrak{F}}(\zeta_{\mathfrak{W}}^S)^i$ for an automorphism S over $\mathfrak{F}$ of $\mathfrak{K}$. It is known that $\zeta_{\mathfrak{W}}^S = \zeta_{\mathfrak{W}}^{s_\mu}$ for some s_μ and $\zeta_{\mathfrak{W}}^{ke+i} = \zeta_{\mathfrak{W}}^{is_\mu}$, (17) holds. Conversely (17) holds for some k if and only if d divides $i(q_{\mathfrak{F}} - 1)$. It is simple then to verify that (16) is an automorphism of $\mathfrak{K} = \mathfrak{W}(P)$ over $\mathfrak{F}$, $P^e = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i$, if and only if $P^e = (P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i)^S$ for S inducing an automorphism over $\mathfrak{F}$ of $\mathfrak{W}$.

4. Normal fields

The most interesting algebraic extensions of a field $\mathfrak{F}$ are the normal fields over $\mathfrak{F}$. For such fields we have the following existence theorem.

THEOREM 5. *There exists a normal field $\mathfrak{K}$ of ramification order e prime to p over a p-adic field $\mathfrak{F}$ if and only if e divides $q_{\mathfrak{K}} - 1$, that is, the maximal unramified subfield of K contains a primitive e^{th} root of unity.*

Our theorem thus states that there exists a normal field $\mathfrak{K}$ of ramification order e prime to p and degree ef over $\mathfrak{K}$ if and only if there are e fields in the set (13) to which all fields with the given e and f are equivalent. We derive it as the consequence[12] of the case $i = 0$ of

THEOREM 6. *Let $\mathfrak{K}$ have degree ef and ramification order e prime to p over F, $K = \mathfrak{W}(P)$, $P^e = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i$. Then $\mathfrak{K}$ is normal over F if and only if e divides both $q_{\mathfrak{F}}^f - 1$ and $i(q_{\mathfrak{F}} - 1)$.*

For $\mathfrak{W}$ is cyclic over $\mathfrak{F}$ and its automorphisms have been seen to carry the minimum function $g(x) = x^e - P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i$ of P over $\mathfrak{W}$ to $g_\mu(x) = x^e - P_{\mathfrak{F}}(\zeta_{\mathfrak{W}}^{q_{\mathfrak{F}}^\mu})^i$. Evidently $\mathfrak{K}$ is normal over $\mathfrak{F}$ if and only if it contains every root of all the polynomials $g_\mu(x)$ in the root field $\mathfrak{Z}$ containing $\mathfrak{K}$ of the product $g_0(x) \cdots g_f(x)$. In particular ζP is a root of $g(x)$ in $\mathfrak{Z}$, where ζ is a primitive e^{th} root of unity. Hence if $\mathfrak{K}$ is normal over $\mathfrak{F}$ the quantity $(\zeta P)P^{-1} = \zeta$ must be in $\mathfrak{K}$. By Lemma 4 this is possible if and only if e divides $q_{\mathfrak{K}} - 1$. If then ζ is in $\mathfrak{K}$ the field $\mathfrak{K}$ will contain all roots of all the $g_\mu(x)$ if and only if $\mathfrak{K}$ contains a root of $g_\mu(x)$ for all μ. Theorem 4 states that this is satisfied if and only if e divides $i(q_{\mathfrak{F}}^\mu - 1)$ for all $\mu = 0, 1, \cdots, e - 1$. But if e divides $i(q_{\mathfrak{F}} - 1)$ it clearly divides $i(q_{\mathfrak{F}}^\mu - 1) = i(q_{\mathfrak{F}} - 1)\,(1 + q_{\mathfrak{F}} + \cdots + q_{\mathfrak{F}}^{\mu-1})$ for all values of μ.

As an almost immediate corollary we have

THEOREM 7. *A completely ramified field of degree e prime to p over a p-adic field $\mathfrak{F}$ is normal over $\mathfrak{F}$ if and only if it, as well as all other completely ramified fields of the same degree over $\mathfrak{F}$, are cyclic over $\mathfrak{F}$. Moreover this occurs when and only when $\mathfrak{F}$ contains a primitive e^{th} root of unity, there are e inequivalent completely ramified fields of degree e over $\mathfrak{F}$.*[13]

For if $\mathfrak{K}$ is normal we apply Theorem 5 with $\mathfrak{F} = \mathfrak{W}$, $f = 1$ to see that e divides $q_{\mathfrak{F}} - 1$, $\mathfrak{F}$ contains a primitive e^{th} root of unity. Every completely

[12] That e necessarily divides $q_{\mathfrak{K}} - 1$ follows explicitly from Theorem 6. The existence of a normal field is given by the case $i = 0$ of that theorem.

[13] See also **[26]**, **[27]** for results of this type.

ramified field of degree e over $\mathfrak{F}$ is then equivalent to one of the fields $\mathfrak{K}_i = \mathfrak{F}(P)$, $P^e = P_{\mathfrak{F}}\zeta_{\mathfrak{F}}^i$ $(i = 0, 1, \cdots, e - 1)$, the correspondence $P \leftrightarrow P^S = \zeta P$ clearly induces an automorphism of order e over $\mathfrak{F}$ of $\mathfrak{K}_i$. Hence every $\mathfrak{K}_i$ of (13) is cyclic over $\mathfrak{F}$.

The result above has the following rather interesting consequence.

THEOREM 8. *Let e be prime to p, Φ be the Euler*[14] *Φ-function of $\Psi = e(q_{\mathfrak{F}}^f - 1)$, ζ_0 be a primitive $(q_{\mathfrak{F}}^{\Phi} - 1)^{\text{th}}$ root of unity, $\mathfrak{W}_0 = \mathfrak{F}(\zeta_0)$. Then every field $\mathfrak{K}$ of ramification order e and degree ef over $\mathfrak{F}$ is equivalent over $\mathfrak{F}$ to a subfield of the normal field*

$$\mathfrak{N} = \mathfrak{W}_0(P) = \mathfrak{W}_0 \times \mathfrak{F}(P) = (\mathfrak{K}, \mathfrak{W}_0), \qquad P^e = P_{\mathfrak{F}}, \tag{18}$$

of ramification order e and degree $e\Phi$ over $\mathfrak{F}$.

For if $\mathfrak{N} = \mathfrak{W}_0(P)$ is defined by $P^e = P_{\mathfrak{F}}$ the field $\mathfrak{N}$ is normal over $\mathfrak{F}$ by Theorem 5 if and only if e divides $q_{\mathfrak{F}}^{\Phi} - 1$. But Φ is defined so that Ψ divides $q_{\mathfrak{F}}^{\Phi} - 1$. Hence $\mathfrak{N}$ is normal over $\mathfrak{F}$. Evidently $\mathfrak{N}$ is the composite of the completely ramified field $\mathfrak{F}(P)$ and the unramified field $\mathfrak{W}_0$, $\mathfrak{N} = \mathfrak{W}_0 \times \mathfrak{F}(P)$. By Lemma 4 $\mathfrak{W}_0$ contains a primitive Ψ^{th} root of unity ζ, ζ^e is a primitive $(q_{\mathfrak{F}}^f - 1)^{\text{th}}$ root of unity $\zeta_{\mathfrak{W}}$ and we may assume that the maximal unramified subfield of $\mathfrak{K}$ is $\mathfrak{W} = \mathfrak{F}(\zeta_{\mathfrak{W}}) \leqq \mathfrak{W}_0$. Now if $P_i = P\zeta^i$ we have P_i in $\mathfrak{N}$, $P_i^e = P_{\mathfrak{F}}\zeta^{ei} = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i$, $\mathfrak{N}$ contains every $\mathfrak{K}_i$ in the sense of equivalence. Clearly $\mathfrak{N}$ is the composite of $\mathfrak{W}_0$ and $\mathfrak{K}$. This proves our theorem.

We shall proceed to determine the automorphism group $\mathfrak{G}$ of any normal field $\mathfrak{K}$ with $(e, p) = 1$. By Theorem 7 $\mathfrak{K} = \mathfrak{W}(P)$ is cyclic over $\mathfrak{W}$, the group of $\mathfrak{K}$ over $\mathfrak{W}$ is $\mathfrak{H} = [T]$. We are taking $\mathfrak{K} = \mathfrak{W}(P)$, $P^e = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i$ and may thus assume that T is induced in $\mathfrak{K}$ by

$$\zeta_{\mathfrak{W}}^T = \zeta_{\mathfrak{W}}, \qquad P^T = \zeta P, \tag{19}$$

where

$$q_{\mathfrak{F}}^f - 1 = er, \qquad \zeta = \zeta_{\mathfrak{W}}^r. \tag{20}$$

The group $\mathfrak{H}$ is a normal divisor of index f of $\mathfrak{G}$ and the quotient group is the cyclic automorphism group

$$\mathfrak{G}/\mathfrak{H} = [S\mathfrak{H}], \tag{21}$$

where S is an automorphism over $\mathfrak{F}$ of $\mathfrak{K}$. Here we may choose S to be any automorphism over $\mathfrak{F}$ of $\mathfrak{K}$ such that $[S\mathfrak{H}]$ induces the group of $\mathfrak{W}$ over $\mathfrak{F}$, that is any S such that

$$\zeta_{\mathfrak{W}} \leftrightarrow (\zeta_{\mathfrak{W}})^S \tag{22}$$

is a generating automorphism of $\mathfrak{W}$ over $\mathfrak{F}$. Thus we take

$$(\zeta_{\mathfrak{W}})^S = (\zeta_{\mathfrak{W}})^{q_{\mathfrak{F}}}. \tag{23}$$

[14] This Φ is the number of integers not greater than and prime to Ψ. We do not use the customary symbol ϕ of the Theory of Numbers, as we shall use the notation ϕ for valuations in a later section.

The choice (23) now implies that

$$\mathfrak{G} = \mathfrak{H} + S\mathfrak{H} + \cdots + S^{f-1}\mathfrak{H}, \tag{24}$$

since the cosets $(S\mathfrak{H})^j = S^j\mathfrak{H}$ are distinct for $j = 0, 1, \cdots, f-1$, $(S\mathfrak{H})^f$ leaves $\zeta_{\mathfrak{W}}$ unaltered and is in $\mathfrak{H}$. Thus there exists a rational integer δ such that

$$S^f = T^\delta. \tag{25}$$

Also $\mathfrak{H}$ is a normal divisor of $\mathfrak{G}$ and

$$TS = ST^\gamma \tag{26}$$

for a rational integer γ. We determine $\mathfrak{G}$ completely in the following

THEOREM 9. *Let $\mathfrak{K}$ be normal of ramification order e prime to p and degree ef over $\mathfrak{F}$ so that $\mathfrak{K} = \mathfrak{W}(P)$, $P^e = P_{\mathfrak{F}}\zeta_{\mathfrak{W}}^i$. Then the group of $\mathfrak{K}$ over $\mathfrak{F}$ consists of the ef distinct automorphisms*

$$S^kT^j \qquad (k = 0, 1, \cdots, f-1, j = 0, 1, \cdots, e-1), \tag{27}$$

where

$$S^f = T^i, \qquad T^e = I, \qquad TS = ST^{q_{\mathfrak{F}}}, \tag{28}$$

and T is given by (19), *S is given by* (23) *and*

$$i(q_{\mathfrak{F}} - 1) = eh, \qquad P^S = \zeta_{\mathfrak{W}}^h P. \tag{29}$$

We have already a partial proof of this result and to complete our proof we need only verify that $\gamma = q_{\mathfrak{F}}$, $\delta = i$ satisfy (26), (25) if (29) holds and that $(\zeta_{\mathfrak{W}}^h P)^e = \zeta_{\mathfrak{W}}^{i(q_{\mathfrak{F}}-1)}\zeta_{\mathfrak{W}}^i P_{\mathfrak{F}} = (\zeta_{\mathfrak{W}}^S)^i P_{\mathfrak{F}}$. Now $P^{TS} = (\zeta_{\mathfrak{W}}^r P)^S = \zeta_{\mathfrak{W}}^{rq_{\mathfrak{F}}+h}P$. Also $P^{T^{q_{\mathfrak{F}}}} = \zeta^{q_{\mathfrak{F}}}P = \zeta_{\mathfrak{W}}^{rq_{\mathfrak{F}}}P$, $P^{ST^{q_{\mathfrak{F}}}} = (\zeta_{\mathfrak{W}}^h P)^{T^{q_{\mathfrak{F}}}} = \zeta_{\mathfrak{W}}^{h+rq_{\mathfrak{F}}}P$. Hence $TS = ST^{q_{\mathfrak{F}}}$ as desired. We compute $P^{S^k} = \zeta_{\mathfrak{W}}^{h(1+q_{\mathfrak{F}}+\cdots+q_{\mathfrak{F}}^{k-1})}P$, and see that $P^{S^{k+1}} = \zeta_{\mathfrak{W}}^{q_{\mathfrak{F}}h(1+q_{\mathfrak{F}}+\cdots+q_{\mathfrak{F}}^{k-1})}\zeta_{\mathfrak{W}}^{q_{\mathfrak{F}}}P = \zeta_{\mathfrak{W}}^{h(1+q_{\mathfrak{F}}+\cdots+q_{\mathfrak{F}}^{k})}P$. Now $h = e^{-1}i(q_{\mathfrak{F}} - 1) = i[e(q_{\mathfrak{F}} - 1)^{-1}]^{-1} = i[(q_{\mathfrak{F}}^f - 1)r^{-1}(q_{\mathfrak{F}} - 1)^{-1}]^{-1} = ir(1 + q_{\mathfrak{F}} + \cdots + q_{\mathfrak{F}}^{f-1})^{-1}$. Hence $P^{S^f} = \zeta_{\mathfrak{W}}^{ir}P = P^{T^i}$, $S^f = T^i$.

In the discussion we shall now make of abelian fields we shall require the following

LEMMA 7. *Let $t = (e, i)$. Then the order of the automorphism S of the group in Theorem 7 is eft^{-1}.*

For if S has order k then $(\mathfrak{H}S)^k = \mathfrak{H}$, k is divisible by the order f of the coset $\mathfrak{H}S$, $k = fl$. Then $S^{fl} = T^{il} = I$. The order of T^i is evidently et^{-1}, et^{-1} divides l. But $S^{eft^{-1}} = T^{et^{-1}} = I$, $k = fl$ divides eft^{-1}, l divides et^{-1}, $l = et^{-1}$, $k = eft^{-1}$ as desired.

5. Abelian fields with e prime to p

Equation (28) implies that $ST = TS$ if and only if $q_{\mathfrak{F}} \equiv 1 \pmod{e}$. But $\mathfrak{K}$ is abelian if and only if $ST = TS$ and we have[15]

[15] The author originally obtained this result for the most interesting case of cyclic fields as follows. Let $\mathfrak{K}$ be cyclic of ramification order e prime to p and degree ef over a

THEOREM 10. *There exists an abelian field of ramification order e prime to p over $\mathfrak{F}$ if and only if all fields of this ramification order are abelian, $\mathfrak{F}$ contains a primitive e^{th} root of unity.*[16]

Thus there are either no abelian fields of ramification order e over $\mathfrak{F}$ or all fields of ramification order e are abelian. Moreover in the latter case for every degree $n = ef$ there is a set of e inequivalent abelian fields of ramification order e and degree n over $\mathfrak{F}$ to one of which every abelian field with the given n and e is equivalent. We also have the elementary consequence.

THEOREM 11. *Let $\mathfrak{F}$ be a p-adic field and n be an integer prime to $p(q_{\mathfrak{F}} - 1)$. Then the only abelian field of degree n over $\mathfrak{F}$ is the cyclic unramified field.*

We continue our study of the structure of abelian fields by proving

THEOREM 12. *Let $\mathfrak{K}$ be an abelian field of degree $n = ef$, ramification order e and a type number i over $\mathfrak{F}$, e be prime to p and define*

$$\delta = (e, f). \tag{30}$$

Then K is the direct product of its maximal unramified subfield $\mathfrak{W}$ and a cyclic completely ramified field of degree e over $\mathfrak{F}$ if and only if δ divides i.

For the congruence

$$tf \equiv i \pmod{e} \tag{31}$$

has a solution t if and only if δ divides i. Let δ divide i, t be a solution of (31) Since $\mathfrak{K}$ is abelian we have $q_{\mathfrak{F}} \equiv 1 \pmod{e}$,

$$\sigma = 1 + q_{\mathfrak{F}} + \cdots + q_{\mathfrak{F}}^{f-1} \equiv f \pmod{e}, \tag{32}$$

so that

$$t\sigma \equiv i \pmod{e}. \tag{33}$$

But then $t\sigma$ is a type number of $\mathfrak{K}$, $\mathfrak{K} = \mathfrak{W}(P)$, $P^e = P_{\mathfrak{F}}(\zeta_{\mathfrak{W}}^{\sigma})^t = P_{\mathfrak{F}}\zeta_{\mathfrak{F}}^{i}$ in $\mathfrak{F}$, $\mathfrak{K} = \mathfrak{W} \times \mathfrak{F}(P)$. Conversely if $\mathfrak{K} = \mathfrak{W} \times \mathfrak{F}(P)$ as above we have $\mathfrak{K} = \mathfrak{W}(P)$, σt is a type number of $\mathfrak{K}$, $\sigma t \equiv i \pmod{e}$, δ must divide i.

The type of abelian field considered above is one extreme, the other of which is considered in

THEOREM 13. *Let $\mathfrak{K}$ be as in the hypothesis of Theorem 12. Then $\mathfrak{K}$ is cyclic over $\mathfrak{F}$ if and only if i is prime to δ.*

For let i be prime to δ, $n = \rho_1^{g_1} \cdots \rho_t^{g_t}$ for positive rational integers g_j, distinct primes ρ_j, and $n = ef$ the degree of $\mathfrak{K}$ over $\mathfrak{F}$. Then

$$e = \rho_1^{\alpha_1} \cdots \rho_t^{\alpha_t}, \qquad f = \rho_1^{\beta_1} \cdots \rho_t^{\beta_t} \tag{34}$$

---

p-adic field $\mathfrak{F}$, so that $\mathfrak{K} = \mathfrak{W}(P)$ as in Theorem 3. The automorphism group of $\mathfrak{K}$ over $\mathfrak{W}$ is the cyclic group $[T]$ where $T = S^f$, S is a generating automorphism of $\mathfrak{K}$ over $\mathfrak{F}$, and clearly $P^T = \omega P$, ω is a primitive e^{th} root of unity. By Lemma 6 $P^S = \zeta_0 P$ for ζ_0 in $\mathfrak{W}$, $P^{S^2} = (\zeta_0^S \zeta_0)P$, $P^T = N_{\mathfrak{W}/\mathfrak{F}}(\zeta_0)P$, so that $\omega = N_{\mathfrak{W}/\mathfrak{F}}(\zeta_0)$ is in $\mathfrak{F}$ as desired.

[16] See also [22] for results of this type.

where the α_j and β_j are rational integers such that

$$(35) \qquad \alpha_j + \beta_j = g_j, \quad \alpha_j \geqq 0, \quad \beta_j \geqq 0 \quad (j = 1, \cdots, t).$$

If $\alpha_j\beta_j > 0$ then ρ_j divides δ and is prime to i, (e, i) is prime to δ_j, $e[(e, i)]^{-1}$ is divisible by $\delta_j^{\alpha_j}$. By Lemma 7 the order of $[S]$ is divisible by $\rho_j^{\alpha_j+\beta_j} = \rho_j^{g_j}$, $[S]$ contains an element S_j of order $\rho_j^{g_j}$. If $\alpha_j = 0$ the integer e is prime to ρ, the order of S is again divisible by $\rho_j^{g_j} = \rho_j^{\beta_j}$, S_j of order $\rho_j^{g_j}$ again exists in $[S]$. Finally if $\beta_j = 0$ the group $[T]$ contains an element S_j of order $\rho_j^{g_j} = \rho_j^{\alpha_j}$. The order of the product $S_0 = S_1 \cdots S_t$ is n, $G = [S_0]$ is cyclic. Conversely let G be cyclic of order n and let ρ be a prime divisor of all the integers e, f, and i. If $G = [S_0]$ then $n = \rho^\gamma n_0$ for $(n_0, \rho) = 1$, $S_1 = S_0^{n_0}$ has order $\rho^\gamma = \rho^{\alpha+\beta}$ where $e = \rho^\alpha e_0$, $f = \rho^\beta f_0$, $(e_0 f_0, \rho) = 1$. The order of S is $\nu = \rho^\lambda \nu_0$, $(\nu_0, \rho) = 1$, the order of S^{n_0} is ρ^λ since n_0 is divisible by ν_0 and is prime to ρ. But ρ divides (e, i), $\lambda < \alpha + \beta$. Thus if $S_0 = S^a T^b$ we have $S_1 = (S^{n_0})^a (T^{n_0})^b$ where T has order ρ^α, $S_1^m = I$ if $m = \rho^{\alpha+\beta-1}$, a contradiction.

As final special case we derive

THEOREM 14. *Let $\mathfrak{K}$ be as in the hypotheses of Theorem 12, and let $i = \Lambda i_0$ where i_0 is prime to ef and Λ divides $\delta = (e, f)$. Then $\mathfrak{K} = \mathfrak{K}_1 \times \mathfrak{K}_2$ where $\mathfrak{K}_2$ is a completely ramified cyclic field of degree Λ over $\mathfrak{F}$, $\mathfrak{K}_1$ is cyclic of degree $ef\Lambda^{-1}$ and ramification order e over $\mathfrak{F}$.*

For $e = \Lambda e_0$, $f = \Lambda f_0$. We use the group of $\mathfrak{K}$ over $\mathfrak{F}$ as determined in Theorem 9 and have $S^f = T^{i_0\Lambda}$, $T_0 = T^{i_0}$ generates the group $[T]$, $T^{i_0\Lambda}$ generates a subgroup of order e_0 of $[T]$. By Lemma 7 the order of S is $fe_0 = ef_0$. We put $T_1 = T_0 S^{-f_0}$ and have $T_1^\Lambda = T^i S^{-f} = I$. The subgroup of $\mathfrak{G}$ generated by S and T_1 contains $T_1 S^{f_0} = T_0$ and hence T, it is the group $\mathfrak{G}$. Its order is at most the product of the order Ef_0 of S by the order of T_1. The order of T_1 divides Λ, the order of $\mathfrak{G}$ is $ef_0\Lambda$, T_1 has order Λ, $\mathfrak{G} = [S] \times [T_1]$. The subfield $\mathfrak{K}_2$ of $\mathfrak{K}$ unaltered by S has intersection $\mathfrak{F}$ with $\mathfrak{W}$, $\mathfrak{K}_2$ is a completely ramified field of degree Λ over $\mathfrak{F}$. The intersection of the subfield $\mathfrak{K}_1$ of $\mathfrak{K}$ unaltered by T_1 with $\mathfrak{W}$ is the set $\mathfrak{W}_0$ of all quantities of $\mathfrak{K}$ unaltered by both T_1 and T. Hence $\mathfrak{W}_0$ is the subfield of $\mathfrak{W}$ unaltered by $T_0 T_1^{-1} = S^{f_0}$, $\mathfrak{W}_0$ is a field of degree f_0 over $\mathfrak{F}$. Now $\mathfrak{K}_1$ is cyclic of degree $f_0 e$ and residue-class degree f_0 over $\mathfrak{F}$, the ramification order of $\mathfrak{K}_1$ over $\mathfrak{F}$ is e.

We shall combine the results above by the use of

THEOREM 15. *Let $\mathfrak{K}$ be an abelian field of degree $n = ef$, ramification order e prime to p and a type number i over $\mathfrak{F}$. Suppose that $n = ab$ for relatively prime a and b so that*

$$e = e_1 e_2, \qquad f = f_1 f_2, \qquad a = e_1 f_1, \qquad b = e_2 f_2$$

for rational integers e_1, e_2, f_1, f_2. Then $\mathfrak{K}$ is the direct product

$$\mathfrak{K} = \mathfrak{L} \times \mathfrak{M}$$

where $\mathfrak{L}$ and $\mathfrak{M}$ are abelian of respective degrees a and b and ramification orders e_1 and e_2 over $\mathfrak{F}$. Moreover if we determine rational integers α and β so that

$$\alpha f_2 \equiv 1 (\mathrm{mod}\ a), \qquad \beta f_1 \equiv 1 (\mathrm{mod}\ b),$$

then $\mathfrak{L}$ has a type number αi and $\mathfrak{M}$ has a type number βi over $\mathfrak{F}$.

For $\mathfrak{K} = \mathfrak{W}(P)$, $P^e = \zeta_{\mathfrak{W}}^i P_{\mathfrak{F}}$, and e divides $q_{\mathfrak{F}} - 1$. Then e divides $q_{\mathfrak{F}}^{f_1} - 1$ and we put

$$\tau = q_{\mathfrak{F}}^f - 1 = (q_{\mathfrak{F}}^{f_1} - 1)\sigma_1, \ \zeta_1 = \zeta_{\mathfrak{W}}^{\sigma}, \tag{36}$$

so that $\mathfrak{W}_1 = \mathfrak{F}(\zeta_1)$ has degree f_1 over $\mathfrak{F}$. Clearly

$$\sigma_1 = 1 + (q_{\mathfrak{F}}^{f_1}) + \cdots + (q_{\mathfrak{F}}^{f_1})^{f_2 - 1} \equiv f_2 (\mathrm{mod}\ e). \tag{37}$$

Hence $\alpha f_2 \equiv \alpha\sigma_1 \equiv 1(\mathrm{mod}\ e)$, $\alpha\sigma_1 \equiv 1(\mathrm{mod}\ e_1)$, and there exists a rational integer γ_1 such that

$$\alpha\sigma_1 = 1 - \gamma_1 e_1 . \tag{38}$$

We put

$$P_1 = P_{\mathfrak{F}}^{\gamma_1} P^{e_2 \alpha \sigma_1}, \tag{39}$$

and have

$$P_1^{e_1} = P_{\mathfrak{F}}^{\gamma_1 e_1} P_{\mathfrak{F}}^{\alpha\sigma_1} (\zeta_{\mathfrak{W}}^{\sigma_1})^i = P_{\mathfrak{F}} \zeta_1^{i\alpha}. \tag{40}$$

The field $\mathfrak{L} = \mathfrak{W}_1(P_1)$ has degree $a = e_1 f_1$, ramification order e_1, and a type number αi over $\mathfrak{F}$, e divides $q_{\mathfrak{F}} - 1$, $\mathfrak{L}$ is abelian. Similarly we obtain an abelian subfield $\mathfrak{M}$ of degree $b = e_2 f_2$, ramification order e_2 and a type number βi over $\mathfrak{F}$. The composite in $\mathfrak{K}$ of its subfields $\mathfrak{L}$ and $\mathfrak{M}$ is their direct product of degree $ab = n$ over $\mathfrak{F}$, $\mathfrak{K} = \mathfrak{L} \times \mathfrak{M}$.

Our final theorem on the structure of p-adic abelian fields is now stated as

THEOREM 16. *Every non-cyclic abelian field of ramification order prime to p over a p-adic field $\mathfrak{F}$ is the direct product of two cyclic fields one of which may be taken to be completely ramified over $\mathfrak{F}$.*

For if $\mathfrak{K}$ is non-cyclic the greatest common divisor Δ of e, f, and i is not unity. Then we may write

$$e = e_0 e_1 , \qquad f = f_0 f_1 , \tag{41}$$

where e_0 is the product of all prime factors of e not dividing Δ, f_0 is the similar product for f. Then $(e_0 f_0, e_1 f_1) = 1$, every prime factor of $e_1 f_1$ divides Δ, that is, divides e_1, f_1, and i. By Theorem 12 $\mathfrak{K} = \mathfrak{L} \times \mathfrak{M}$ where $\mathfrak{L}$ is abelian of degree $e_1 f_1$, ramification order e_1 and a type number αi over $\mathfrak{F}$, $\alpha f_0 \equiv 1(\mathrm{mod}\ e_1 f_1)$, α is prime to $e_1 f_1$. Hence the greatest common divisor of e_1, f_1, and αi is Δ. Similarly $\mathfrak{M}$ is abelian of degree $e_0 f_0$ and ramification order e_0 over $\mathfrak{F}$, β is prime to $e_0 f_0$, a type number of $\mathfrak{M}$ is βi prime to $e_0 f_0$, $\mathfrak{M}$ is cyclic by Theorem 13. There remains the discussion of the structure of $\mathfrak{L}$.

If ρ is any prime factor of e_1f_1 we have determined e_1 and f_1 so that ρ divides e_1, f_1, and i. We put the prime factors of e_1f_1 into two classes. Let the first class consist of those primes ρ for which the highest power of ρ dividing $\delta = (e, f)$ divides i, and the second class all remaining primes. Then if ρ is in the second class and $e_1 = \rho^{\lambda_1}e_2$, $f_1 = \rho^{\lambda_2}f_2$, $i = \rho^{\lambda_3}i_2$, we have chosen ρ so that the minimum of λ_1 and λ_2 is greater than λ_3. We may then write

$$e_1 = e_3e_4, \qquad f_1 = f_3f_4, \tag{42}$$

where

$$(e_3f_3, e_4f_4) = 1, \tag{43}$$

e_3f_3 is the product of primes of the first class, e_4f_4 of primes of the second class. Clearly (e_3, f_3) divides i, $\beta i = \Lambda i_0$ where i_0 is prime to e_4f_4 and Λ divides (e_4, f_4). Then by Theorem 14

$$\mathfrak{L} = \mathfrak{L}_3 \times \mathfrak{L}_4 \tag{44}$$

where $\mathfrak{L}_j$ has degree e_jf_j, ramification order e_j, and a type number $\alpha_j i$ over $\mathfrak{F}$. By Theorem 12 $\mathfrak{L}_3 = \mathfrak{W}_3 \times \mathfrak{K}_3$, $\mathfrak{W}_3$ is unramified of degree f_3 over $\mathfrak{F}$, $\mathfrak{K}_3$ is completely ramified of degree e_3 over $\mathfrak{F}$. By Theorem 13, $\mathfrak{L}_4 = \mathfrak{L}_5 \times \mathfrak{K}_4$ over $\mathfrak{F}$, $\mathfrak{K}_4$ is completely ramified of degree Λ over $\mathfrak{F}$, $\mathfrak{L}_5$ has ramification order e_4 and residue-class degree $f_4\Lambda^{-1}$ over $\mathfrak{F}$. Thus

$$\mathfrak{K} = (\mathfrak{M} \times \mathfrak{W}_3 \times \mathfrak{L}_3) \times (\mathfrak{L}_5 \times \mathfrak{K}_4) \tag{45}$$

where $\mathfrak{M} \times \mathfrak{W}_3 \times \mathfrak{L}_3$ is cyclic over $\mathfrak{F}$, $\mathfrak{L}_5 \times \mathfrak{K}_4$ is a cyclic completely ramified field. This proves our theorem.

In closing we note

THEOREM 17. *Let e and f be rational integers, $(e, p) = 1$, and write $\delta = (e, f)$, $e = \delta\epsilon$. Then every cyclic field of degree ef and ramification order e over a p-adic field $\mathfrak{F}$ is equivalent to one of a set of $\epsilon\Phi(\delta)$ such fields inequivalent over $\mathfrak{F}$, where $\Phi(\delta)$ is the Euler Φ-function.*

For we have seen that every abelian field is equivalent to one and only one of the fields with type numbers $i = 0, 1, \cdots, e - 1$ and these fields are cyclic if and only if i is prime to δ. There are $\Phi(\delta)$ numbers not greater than δ and prime to δ and if a is one of these then $a, a + \delta, a + 2\delta, \cdots, a + (\epsilon - 1)\delta$ are complete set of integers in the set $0, 1, \cdots, e - 1$ which are congruent to a modulo δ. Evidently every desired integer i is congruent to an a modulo δ, there are $\epsilon\Phi(\delta)$ such integers.

6. Structure of arbitrary fields

Our first results on the structure of fields whose ramification order is arbitrary is the rather surprising

THEOREM 18. *Let $\mathfrak{K}$ be a field of degree ef and ramification order e over a p-adic field $\mathfrak{F}$ and let*

$$e = p^{\nu}e_0, \qquad (e_0, p) = 1. \tag{46}$$

Then $\mathfrak{K}$ is completely ramified of degree p^{ν} over a subfield $\mathfrak{L}$ of degree e_0f and ramification order e_0 over $\mathfrak{F}$.

For, by Theorem 1, $\mathfrak{K} = \mathfrak{W}(P)$, $P^e = UP_{\mathfrak{F}}$, for a unit U of $\mathfrak{K}$. We write $U = \zeta_{\mathfrak{W}}^i U_0$ where U_0 is a principal unit of $\mathfrak{K}$. By Lemma 5 $U_0 = V^{e_0}$ for a principal unit V of $\mathfrak{K}$. Pub $\mathfrak{L} = \mathfrak{W}(R)$, $R = V^{-1}P^{p}$ and have $R^{e_0} = V^{e_0}P^e = U_0^{-1}\zeta_{\mathfrak{W}}^i U_0 P = \zeta_{\mathfrak{W}}^i P_{\mathfrak{F}}$, $\mathfrak{L}$ has degree e_0f and ramification order e_0 over $\mathfrak{F}$, $\mathfrak{K} = \mathfrak{L}(P)$, $P^{p^{\nu}} = VR$, $\mathfrak{K}$ is completely ramified of degree p^{ν} over $\mathfrak{L}$. This completes our proof.

If $\mathfrak{W}$ is the maximal unramified subfield over $\mathfrak{F}$ of a field $\mathfrak{K}$ over $\mathfrak{F}$ and $\mathfrak{W}_0$ is any unramified extension of $\mathfrak{W}$ the composites

$$\mathfrak{K}_0 = (\mathfrak{K}, \mathfrak{W}_0), \qquad \mathfrak{L}_0 = (\mathfrak{L}, \mathfrak{W}_0) \tag{47}$$

over $\mathfrak{F}$ are clearly the direct products

$$\mathfrak{K}_0 = \mathfrak{K} \times \mathfrak{W}_0, \qquad \mathfrak{L}_0 = \mathfrak{L} \times \mathfrak{W}_0 \tag{48}$$

over $\mathfrak{W}$. Hence the ramification order of $\mathfrak{K}_0$ over $\mathfrak{F}$ coincides with that of $\mathfrak{K}$ over $\mathfrak{F}$, $\mathfrak{L}_0$ is the field of Theorem 18 defined for $\mathfrak{K}_0$. We may then prove

THEOREM 19. *Every field $\mathfrak{K}$ of finite degree over a p-adic field is contained in a field $\mathfrak{K}_0$ of the same ramification order as $\mathfrak{K}$ over $\mathfrak{F}$ and such that the field $\mathfrak{L}_0$ defined in Theorem 18 for $\mathfrak{K}_0$ is normal over $\mathfrak{F}$. Moreover if $\mathfrak{K}$ is normal over $\mathfrak{F}$ the field $\mathfrak{K}_0$ may be taken to be normal over $\mathfrak{F}$.*

For we choose $\mathfrak{W}_0$ as in Theorem 8 and have $L_0 = \mathfrak{L} \times \mathfrak{W}_0$ normal over $\mathfrak{F}$. The composite of a normal field $\mathfrak{K}$ over $\mathfrak{F}$ and a cyclic field $\mathfrak{W}_0$ over $\mathfrak{F}$ is normal over $\mathfrak{F}$.

We use the results obtained thus far to obtain a new proof of Hensel's[17]

THEOREM 20. *The automorphism group of a normal field $\mathfrak{K}$ over a p-adic field $\mathfrak{F}$ is a solvable group.*

For we use the field $\mathfrak{K}_0$ of Theorem 19 to see that $\mathfrak{K}$ is contained in $\mathfrak{K}_0 \geqq \mathfrak{L}_0 \geqq \mathfrak{W}_0 \geqq \mathfrak{F}$, where $\mathfrak{K}_0$ is normal of degree p^{ν} over $\mathfrak{L}_0$, $\mathfrak{L}_0 = \mathfrak{W}_0(P)$ such that $P^{e_0} = P_{\mathfrak{W}_0}$ in $\mathfrak{W}_0$, $\mathfrak{L}_0$ is normal over $\mathfrak{W}_0$, $\mathfrak{W}_0$ is cyclic over $\mathfrak{F}$. The group of $\mathfrak{K}_0$ over $\mathfrak{L}_0$ is a p-group and solvable, $\mathfrak{K}_0$ is metacyclic over $\mathfrak{L}_0$. The field $\mathfrak{L}_0$ is completely ramified and is cyclic over $\mathfrak{W}_0$. Hence $\mathfrak{K}_0$ is metacyclic over $\mathfrak{F}$, the group $\mathfrak{G}_0$ of $\mathfrak{K}_0$ is solvable. The normal subfield $\mathfrak{K}$ of $\mathfrak{K}_0$ corresponds to a subgroup $\mathfrak{H}$ of $\mathfrak{G}_0$, the group of $\mathfrak{K}$ over $\mathfrak{F}$ is $\mathfrak{G} = \mathfrak{G}_0/\mathfrak{H}$ which is homomorphic to $\mathfrak{G}_0$ and hence is solvable.

7. Finiteness of the number of fields

If ϵ is the ramification order of a p-adic field $\mathfrak{F}$ over $\mathfrak{R}_p$ we define

$$\gamma = \epsilon + \psi, \tag{49}$$

where ψ is the greatest integer in the rational fraction $\epsilon(p-1)^{-1}$. We let

[17] See [16] in this connection.

$a_1, \cdots, a_\gamma$ range independently over $0, 1, \zeta_{\mathfrak{F}}, \cdots, \zeta_{\mathfrak{F}}^{q_{\mathfrak{F}}-2}$ so that there are $q_{\mathfrak{F}}$ principal units in the set of principal units of the form

$$V = 1 + a_1 P_{\mathfrak{F}} + a_2 P_{\mathfrak{F}}^2 + \cdots + a_\gamma P_{\mathfrak{F}}^\gamma, \tag{50}$$

every principal unit of $\mathfrak{F}$ is congruent modulo $p_{\mathfrak{F}}^{\gamma+1}$ to one of the units (50). Then we may prove

LEMMA 8. *Every principal unit in $\mathfrak{F}$ has the form $U = U_0^p V$, where U_0 is a principal unit of $\mathfrak{F}$ and V is one of the $q_{\mathfrak{F}}^\gamma$ quantities* (50).

For every principal unit has the form $U = V + P_{\mathfrak{F}}^{\gamma+1} B$ where B is an integer of $\mathfrak{F}$. Then $E = UV^{-1} = 1 + P_{\mathfrak{F}}^{\gamma+1} D$ with $D = BV^{-1}$ an integer of $\mathfrak{F}$. The congruence $y^p \equiv E (\mathrm{mod}\ \mathfrak{p}_{\mathfrak{F}}^{\gamma+1})$ has the solution $y = 1$. By a known result[18] there exists an integer U_0 in $\mathfrak{F}$ such that $U_0^p = E$. Now $U_0^p \equiv 1 (\mathrm{mod}\ \mathfrak{p}_{\mathfrak{F}}^{\gamma+1})$, $(p) = \mathfrak{p}_{\mathfrak{F}}^{e}$, $(U_0 - 1)^p \equiv U_0^p \equiv 0 (\mathrm{mod}\ \mathfrak{p}_{\mathfrak{F}}^{e})$, $U_0 \equiv 1 (\mathrm{mod}\ \mathfrak{p}_{\mathfrak{F}})$, U_0 is a principal unit of $\mathfrak{F}$, $U = VU_0^p$ as desired.

We shall use the result above in a discussion of the finiteness[19] of the number of p-adic fields of a given degree. Note that every subfield of a normal field $\mathfrak{N}$ corresponds uniquely to a subgroup of its automorphism group. This group is a finite group, it has only a finite number of subgroups, $\mathfrak{N}$ has only a finite number of subfields. This is evidently also true of any subfield of $\mathfrak{N}$ and hence every field of finite degree over $\mathfrak{F}$ has only a finite number of subfields.

We now let $\mathfrak{K}$ be cyclic of degree p over a p-adic field $\mathfrak{F}$, ζ be a primitive p^{th} root of unity, $\mathfrak{F}_0$ be the cyclic field $\mathfrak{F}(\zeta)$. Then it is well known that the abelian field $\mathfrak{K}_0 = \mathfrak{K}(\zeta) = \mathfrak{F}_0(\xi)$ where $\xi^p = a$ in $\mathfrak{F}_0$. But $a = P_{F_0}^\lambda \zeta_{F_0}^\mu U$ for a principal unit U of $\mathfrak{F}_0$ and Lemma 8 states that $U = U_0^p V$ for V given by (50). We write $\lambda = \lambda_0 p + r$ for $0 \leqq r < p$, use $\zeta_{\mathfrak{F}_0} = (\zeta_{\mathfrak{F}_0}^k)^p$ for $k = q_{\mathfrak{F}_0} p^{-1}$ and have $a = b^p P_{\mathfrak{F}_0}^r V$. Then $\mathfrak{K}_0 = \mathfrak{F}_0(\eta)$ where $\eta = b^{-1}\xi$, $\eta^p = P_{\mathfrak{F}_0}^r V$ is one of a finite number of quantities of F_0. Hence $\mathfrak{K}$ is equivalent to a field contained as a subfield of one of a finite number of abelian fields over $\mathfrak{F}$.

The result just proved combined with Theorem 16 states that there are only finitely many inequivalent cyclic fields of any prime degree ρ over any p-adic field $\mathfrak{F}$. If $\mathfrak{K}$ is normal of degree n over $\mathfrak{F}$ Theorem 20 implies that $\mathfrak{K} = \mathfrak{K}_t > \mathfrak{K}_{t-1} > \cdots > \mathfrak{K}_1 > \mathfrak{K}_0 = \mathfrak{F}$ where $\mathfrak{K}_1$ is cyclic of prime degree ρ_i over $\mathfrak{K}_{i-1} (i = 1, \cdots, t)$, $n = \rho_1 \cdots \rho_t$, $\mathfrak{K}_i$ is equivalent to one of a finite number of inequivalent fields of degree ρ_i over $\mathfrak{K}_{i-1}$. But then there are only a finite number of such fields $\mathfrak{K}$ in the sense of equivalence. Our earlier argument then implies

THEOREM 21. *The number of inequivalent fields of a fixed degree over a p-adic field $\mathfrak{F}$ is finite.*

8. The Grunwald existence theorem

The very important theorem of Grunwald on the existence of abelian fields with prescribed local properties has never been stated, in the literature, in a

[18] See Theorem 1 of [**29**].

[19] See footnote 5.

form most convenient for applications to the theory of rational division algebras which motivated its discovery. We shall analyze the theorem here and obtain an equivalent form free of the terminology of class field theory and much more useful for the study of algebras.

We let $\mathfrak{F}^*$ be the set of all non-zero quantities α of an algebraic number field $\mathfrak{F}$, $\mathfrak{G}$ be a finite abelian group of order n, and $\phi_1, \cdots, \phi_r$ a finite number of non-archimedean valuations of $\mathfrak{F}$. To each ϕ_j there corresponds a unique prime ideal $\mathfrak{p}_j$ of $\mathfrak{F}$ defining ϕ_j. For each j we assume the existence of a function $\chi_j(\alpha)$ on $\mathfrak{F}^*$ to $\mathfrak{G}$ such that $\chi_j(\alpha_1\alpha_2) = \chi_j(\alpha_1)\chi_j(\alpha_2)$. We also assume that there exists a power, and hence a least power $\mathfrak{C}_j$, of $\mathfrak{p}_j$ such that if $\alpha \equiv 1 \pmod{\mathfrak{C}_j}$ then $\chi_j(\alpha) = I$. The Grunwald theorem states[20] that then there exists a field $\mathfrak{K}$ of degree n and automorphism group $\mathfrak{G}$ over $\mathfrak{F}$, such that the norm residue symbol

$$\left(\frac{\alpha, \mathfrak{K}}{\mathfrak{p}_j}\right) = \chi_j(\alpha) \tag{51}$$

for every α of $\mathfrak{F}^*$. Moreover if $\mathfrak{F}^{(k)}$ ranges over all real conjugates of $\mathfrak{F}$ and n is even we may prescribe the corresponding conjugates $\mathfrak{K}^{(k)}$ over $\mathfrak{F}^{(k)}$ to be either real or imaginary as desired.

Let now $\mathfrak{F}_{\phi_j}$ be the derived field of $\mathfrak{F}$ with respect to ϕ_j, $\mathfrak{Z}_j$ be any abelian field over $\mathfrak{F}_{\phi_j}$ with automorphism group Γ_j equivalent to a subgroup $\mathfrak{G}_j$ of $\mathfrak{G}$. The norm-residue symbol

$$\left(\frac{\alpha, \mathfrak{Z}_j}{\mathfrak{p}_j}\right) \tag{52}$$

is defined in the local class field theory as a function on the set $\mathfrak{F}^*_{\phi_j}$ of all non-zero quantities of $\mathfrak{F}_{\phi_j}$ to Γ_j. If α is in $\mathfrak{F}^*$ we define $\chi_j(\alpha)$ to be the element of $\mathfrak{G}_j$ corresponding to (52) under the given equivalence of Γ_j and $\mathfrak{G}_j$. Then it is known in the local class field theory that $\chi_j(\alpha)$ has the properties in the hypothesis above of Grunwald's theorem and we choose it to be the function of that theorem for $j = 1, \cdots, r$. But $\mathfrak{Z}_j$ is uniquely determined in the sense of equivalence as *the local class-field corresponding to the class group H of all quantities α of $\mathfrak{F}^*_{\phi_j}$ such that* (52) *is the identity element of* Γ_j. Since the composite $\mathfrak{K}_j$ of $\mathfrak{K}$ and $\mathfrak{F}_{\phi_j}$ corresponds to the same class group $\mathfrak{H}$ as does $\mathfrak{Z}_j$ it is equivalent over $\mathfrak{F}_j$ to $\mathfrak{Z}_j$. If ϕ is any archimedean valuation of $\mathfrak{F}$ and $\mathfrak{Z}$ is a field over $\mathfrak{F}_\phi$ such that the group of $\mathfrak{Z}$ over $\mathfrak{F}_\phi$ is equivalent to a subgroup of $\mathfrak{G}$, then either $\mathfrak{Z} = \mathfrak{F}_\phi$ or $\mathfrak{G}$ has even order and $\mathfrak{Z}$ is quadratic over $\mathfrak{F}_\phi$. The latter case can occur only when $\mathfrak{F}_\phi$ is real. Grunwald's theorem then states that we may take the composite of $\mathfrak{K}$ and $\mathfrak{F}_\phi$ to be equivalent to $\mathfrak{Z}_\phi$. We now state our result as

THEOREM 22. *Let $\mathfrak{F}$ be an algebraic number field, $\phi_1, \cdots, \phi_r$ be any finite number of valuations of $\mathfrak{F}$, $\mathfrak{G}$ be an abelian group, $\mathfrak{Z}_j$ be a field over the derived field $\mathfrak{F}_{\phi_j}$ of $\mathfrak{F}$ with automorphism group equivalent to a subgroup of $\mathfrak{G}$. Then there*

[20] See [7].

exists a field $\mathfrak{K}$ *with automorphism group* $\mathfrak{G}$ *over* $\mathfrak{F}$ *such that the composite of* $\mathfrak{K}$ *and* $\mathfrak{F}_{\phi_j}$ *is equivalent over* $\mathfrak{F}_{\phi_j}$ *to* $\mathfrak{Z}_j$.

That there exist infinitely many fields $\mathfrak{K}$ follows from Grunwald's proof. However we may better regard this result as a corollary of our theorem which we state as the

COROLLARY. *There exist infinitely many fields* $\mathfrak{K}$ *inequivalent over* $\mathfrak{F}$ *and with the property of Theorem* 22.

For if t is any integer and ρ is any prime divisor of the order of the group $\mathfrak{G}$ there exist non-archimedean valuations $\phi_{r+1}, \cdots, \phi_{r+t}$ all distinct from the $\phi_1, \cdots, \phi_r$ and fields $\mathfrak{W}_k$ unramified of degree ρ over $\mathfrak{F}_{\phi_{r+k}}$. We apply Theorem 20 to obtain an abelian field $\mathfrak{K}_k$ such that $(\mathfrak{K}_k, \mathfrak{F}_\phi) = \mathfrak{W}_k$, $(\mathfrak{K}_k, \mathfrak{F}_{\phi_{r+i}}) = \mathfrak{F}_{\phi_{r+i}}$ for $i \neq k$. Evidently the fields $\mathfrak{K}_k$ are inequivalent over $\mathfrak{F}$ for distinct values of k.

Our Theorem 22 indicates that the Grunwald theorem should be regarded merely as a result which allows us to obtain a field over $\mathfrak{F}$ with a prescribed abelian group and prescribed composites with a finite number of derived fields of $\mathfrak{F}$. Its hypotheses should not contain any conditions on the existence of such composites but this should be left as a problem of the structure and existence of algebraic extensions of p-adic fields.

9. Splitting fields of rational division algebras

If $\mathfrak{D}$ is any normal division algebra of degree n over an algebraic number field $\mathfrak{K}$ and $\mathfrak{Z}$ is any cyclic field over $\mathfrak{K}$ then it is known[21] that $\mathfrak{Z}$ splits $\mathfrak{D}$ if and only if the degree over $\mathfrak{K}_\phi$ of the composite of $\mathfrak{Z}$ and $\mathfrak{K}_\phi$ is a multiple of the index of $\mathfrak{D} \times \mathfrak{K}_\phi$ for every valuation ϕ of $\mathfrak{K}$. Our form of the Grunwald theorem then implies

THEOREM 23. *Let* $\mathfrak{D}$ *be a normal division algebra of degree* n *over an algebraic number field* $\mathfrak{K}$ *of finite degree over* $\mathfrak{F}$, *and let* $\phi_1, \cdots, \phi_r$ *be all the non-archimedean valuations of* $\mathfrak{K}$ *for which the index of* $\mathfrak{D} \times \mathfrak{K}_{\phi_j}$ *is* $m_j > 1$. *Then there exists a cyclic field* $\mathfrak{Z}$ *of degree* n *over* $\mathfrak{F}$ *whose composite with* $\mathfrak{K}$ *splits* $\mathfrak{D}$ *if and only if there exist cyclic fields* $\mathfrak{Z}_j$ *of degree* ν_j *over* $\mathfrak{F}$ *such that* ν_j *divides* n, *the composite of* $\mathfrak{Z}_j$ *and* $\mathfrak{K}_{\phi_j}$ *has degree a multiple of* m_j *over* $\mathfrak{K}_{\phi_j}$, $j = 1, \cdots, r$.

For, by Theorem 22, there exists a cyclic field $\mathfrak{Z}$ of degree n over $\mathfrak{F}$ such that the composite of $\mathfrak{Z}$ and $\mathfrak{F}_{\phi_j}$ is equivalent to $\mathfrak{Z}_j$, the composite of $\mathfrak{Z}$ and $\mathfrak{F}_\phi$ is algebraically closed for every archimedean valuation ϕ of $\mathfrak{F}$. Let $\mathfrak{Z}_\mathfrak{K}$ be the composite of $\mathfrak{Z}$ and $\mathfrak{K}$. Evidently the degree of $(\mathfrak{Z}_\mathfrak{K}, \mathfrak{K}_\phi)$ is a multiple of the index of $\mathfrak{D} \times \mathfrak{K}_\phi$ for every archimedean valuation ϕ and every non-archimedean $\phi \neq \phi_j$. Now the degree of $(\mathfrak{Z}_\mathfrak{K}, \mathfrak{K}_{\phi_j}) = (\mathfrak{Z}, \mathfrak{K}_{\phi_j}) = (\mathfrak{Z}_j, \mathfrak{K}_{\phi_j})$ is a multiple of m_j as desired.

Let us discuss some consequences of the condition of our theorem. We observe first that we have

THEOREM 24. *Let* $\mathfrak{D}$ *be a normal division algebra of degree* n *over its centrum*

---

21 Cf. [4], Chapter IX.

$\mathfrak{K}$ which has degree q prime to n over an algebraic number field $\mathfrak{F}$. Then there exists a cyclic field $\mathfrak{Z}$ of degree n over $\mathfrak{F}$ such that the composite of $\mathfrak{Z}$ and $\mathfrak{K}$ splits $\mathfrak{D}$.

For we choose all the integers ν_j of Theorem 23 to have the value n. The degree of $\mathfrak{K}\phi_j$ over $\mathfrak{F}_{\phi_j}$ is a divisor of q and prime to n, the composite $(\mathfrak{Z}_i, \mathfrak{K}_{\phi_j}) = \mathfrak{Z}_j \times \mathfrak{K}_{\phi_j}$ has degree n over $\mathfrak{K}_{\phi_j}$, m_j divides ν_j as desired.

We next prove

THEOREM 25. *Let $\mathfrak{D}$ be a normal division algebra of degree $n = 2\nu$ over a field $\mathfrak{K}$ of degree 2μ over an algebraic number field $\mathfrak{F}$ and let μ and ν be relatively prime odd integers. Then there exists a cyclic field $\mathfrak{Z}$ of degree n over $\mathfrak{F}$ such that the composite of $\mathfrak{Z}$ and $\mathfrak{K}$ splits $\mathfrak{D}$.*

For if the maximal unramified subfield $\mathfrak{W}_j$ of $\mathfrak{K}_{\phi_j}$ has odd degree over $\mathfrak{F}_{\phi_j}$ we choose the cyclic field $\mathfrak{Z}_j$ of Theorem 23 to be unramified of degree n over $\mathfrak{F}_{\phi_j}$ of the composite of $\mathfrak{Z}_j$ and $\mathfrak{K}_{\phi_j}$ is their direct product and has degree n over $\mathfrak{K}_{\phi_j}$. If, however, $\mathfrak{W}_j$ has even degree the composite $\mathfrak{Y}_j$ of $\mathfrak{W}_j$ and the $\mathfrak{F}_{\phi_j}(R_j)$, $R_j^2 = P_j$ the prime quantity of $\mathfrak{F}_{\phi_j}$, is the direct product $\mathfrak{W}_j(R_j)$ of degree two over $\mathfrak{W}_j$. The degree of $\mathfrak{K}_{\phi_j}$ over $\mathfrak{W}_j$ divides μ and is odd, the composite $\mathfrak{K}_{\phi_j}(R_j)$ of $\mathfrak{Y}_j$ and $\mathfrak{K}_{\phi_j}$ has degree two over $\mathfrak{K}_{\phi_j}$. We now let $\mathfrak{Z}_j$ be the composite of $\mathfrak{F}_{\phi_j}(R_j)$ and the unramified field of degree ν over $\mathfrak{F}_{\phi_j}$, this latter field has degree over $\mathfrak{F}_{\phi_j}$ prime to the degree two of $\mathfrak{K}_{\phi_j}(R_j)$, the composite of $\mathfrak{Z}_j$ and $\mathfrak{K}_{\phi_j}$ has degree n over $\mathfrak{K}_{\phi_j}$. Hence $(\mathfrak{Z}, \mathfrak{K})$ splits $\mathfrak{D}$.

The result above is a generalization of the case where $\mathfrak{D}$ is a generalized quaternion division algebra over a quadratic extension of the field $\mathfrak{R}$ of all rational numbers. As we indicated in our introduction it is this case that inspired the present article.

The field $\mathfrak{Z}$ may not exist in the remaining types of degrees n and q, and we prove the following existence theorem.

THEOREM 26. *Let n and q be rational integers which are neither relatively prime nor the doubles of odd relatively prime integers. Then there exists a field $\mathfrak{K}$ of degree q over the rational number field $\mathfrak{R}$ and a normal division algebra of degree n over $\mathfrak{K}$ such that $\mathfrak{D}$ is not split by the composite of $\mathfrak{K}$ and any cyclic field of degree n over $\mathfrak{R}$.*

For let ρ be a prime dividing both n and q, $n = \rho^{\nu} n_0$ for $(n_0, \rho) = 1$. Then if $\mathfrak{Z}$ is cyclic of degree n over $\mathfrak{R}$ the field $(\mathfrak{Z}, \mathfrak{K})$ splits $\mathfrak{D}$ if and only if $(\mathfrak{Z}_\rho, \mathfrak{K})$ splits $\mathfrak{D}_\rho$, where $\mathfrak{Z}_\rho$ is the subfield of degree ρ^{ν} over $\mathfrak{K}$ of $\mathfrak{Z}$, and $\mathfrak{D}_\rho$ is the direct factor of $\mathfrak{D}$ of degree ρ^{ν} over $\mathfrak{K}$. Hence we may assume, without loss of generality, that $n = \rho^{\nu}$ where ρ is a rational prime dividing q, and that if $\rho = 2$ then either $\nu > 1$ or q is divisible by four. We now define a rational prime p. If ρ is odd we let p be a prime of the form $p = \rho t + 2$ so that $p - 1 = \rho t + 1$ is prime to n. If $\rho = 2$ but $\nu > 1$ we let $p = 2^{\nu} t + 3$ be a prime, and have $p - 1 = 2(2^{\nu-1} t + 1)$ not divisible by 4. In either case Theorem 8 implies that the only cyclic field of degree n over $\mathfrak{R}_p$ is the unramified field.

We shall construct an abelian field $\mathfrak{K}$ of degree q over $\mathfrak{R}$ and use the notation $\mathfrak{p}$ for a prime ideal divisor of p in $\mathfrak{K}$, ϕ for the corresponding valuation. Apply Theorem 23 to construct a field $\mathfrak{K}$ which is cyclic of degree q over $\mathfrak{R}$ and such

that $\mathfrak{K}_\phi = (\mathfrak{K}, \mathfrak{R}_p)$ is unramified over $\mathfrak{R}_p$. Then the composite of $\mathfrak{K}_\phi$ and any cyclic field of degree n over $\mathfrak{R}_p$ cannot have degree n over $\mathfrak{K}_\phi$.

Let next $\rho = n = 2$, $q = 4r$ for an integer r. We let p by any odd prime, $\mathfrak{K}$ be an abelian field of degree q over $\mathfrak{R}$ chosen so that $\mathfrak{K}_\phi = (\mathfrak{K}, \mathfrak{R}_p)$ is the composite of an unramified field of degree $2r$ over $\mathfrak{R}_p$ and the field $\mathfrak{R}_p(P)$, $P^2 = p$. Then $\mathfrak{K}_\phi$ contains a primitive $(p^2 - 1)^{\text{th}}$ root of unity ζ_0, $p^2 - 1 = (p-1)(p+1) = 2\lambda(p-1)$, $(\zeta_0^\lambda)^2 = \zeta$ is the primitive $(p-1)^{\text{th}}$ root of unity in $\mathfrak{R}_p$. Hence $\mathfrak{K}_\phi$ contains $\zeta_0^\lambda P$ such that $(\zeta_0^\lambda P)^2 = \zeta p$. By Theorem 3 $\mathfrak{K}_\phi$ contains all quadratic fields over $\mathfrak{R}_p$.

We have made the argument above to prove the existence of a field $\mathfrak{K}$ of degree q over $\mathfrak{R}$ and a valuation ϕ of $\mathfrak{K}$ such that the composite of $\mathfrak{K}_\phi$ and any cyclic field $\mathfrak{Z}$ of degree n over $\mathfrak{R}$ cannot have degree n over $\mathfrak{K}_\phi$. By Theorem 22 there exists a field $\mathfrak{Y}$ which is cyclic of degree n over $\mathfrak{K}$ such that $(\mathfrak{Y}, \mathfrak{K}_\phi)$ is unramified of degree n over $\mathfrak{K}_\phi$. We may choose a quantity γ in $\mathfrak{K}$ which has order one with respect to ϕ. Then $\mathfrak{D} = (\mathfrak{Y}, S, \gamma)$ is a cyclic algebra such that $\mathfrak{D} \times \mathfrak{K}_\phi$ is a division algebra. Evidently $(\mathfrak{Z}, \mathfrak{K}_\phi)$ cannot split $\mathfrak{D} \times \mathfrak{K}_\phi$, $(\mathfrak{Z}, \mathfrak{K})$ cannot split $\mathfrak{D}$.

We close our results with a proof of

THEOREM 27. *Let $\mathfrak{K}$ be an algebraic number field of degree q over $\mathfrak{R}$, $\mathfrak{D}$ be a normal division algebra of degree n over $\mathfrak{K}$. Then there exists a cyclic field $\mathfrak{Z}$ of degree nq over $\mathfrak{R}$ such that the composite of $\mathfrak{Z}$ and $\mathfrak{K}$ splits $\mathfrak{D}$.*

For let $\phi_1, \cdots, \phi_r$ be the non-archimedean valuations of $\mathfrak{K}$ such that the index of $\mathfrak{D} \times \mathfrak{K}_{\phi_j}$ is $m_j > 1$. We let $\mathfrak{R}_{\phi_j}$ be the corresponding derived fields of $\mathfrak{R}$ contained in $\mathfrak{K}_{\phi_j}$ and choose a cyclic field $\mathfrak{Z}$ of degree nq over $\mathfrak{R}$ such that $(\mathfrak{Z}, \mathfrak{R}_{\phi_j})$ is unramified of degree nq over $\mathfrak{R}_{\phi_j}$ and the fields $(\mathfrak{Z}, \mathfrak{R}_\phi)$ are algebraically closed for every archimedean valuation ϕ of $\mathfrak{R}$. Then the degree of $(\mathfrak{Z}, \mathfrak{K}_{\phi_j})$ is nqr_j^{-1} where r_j is the degree over $\mathfrak{R}_{\phi_j}$ of the intersection of $\mathfrak{K}_{\phi_j}$ and $\mathfrak{Z}$. Hence r_j divides q, nqr_j^{-1} is a multiple of n and of m_j. It follows that $(\mathfrak{Z}, \mathfrak{K})$ splits $\mathfrak{D}$.

THE UNIVERSITY OF CHICAGO.

BIBLIOGRAPHY

ALBERT, A. A. [1]: *A construction of all non-commutative rational division algebras of order* 8, these Annals, v. 31 (1930), pp. 567–76; [2]: *Division algebras over an algebraic field*, A. M. S. Bulletin, v. 37 (1931), pp. 777–84; [3]: *Modern Higher Algebra*, Chicago, 1937; [4]: *Structure of Algebras*, A. M. S. Colloquium Publications, v. 24, New York, 1939.

CHEVALLEY, C. [5]: *La théorie du symbole de restes normiques*, J. für Math., v. 169 (1932), pp. 141–57; [6]: *Sur la théorie du corps de classes dans les corps finis et les corps locaux*, Tokyo J., v. 2 (1933), pp. 365–474.

GRUNWALD, W. [7]: *Ein allgemeines Existenztheorem für algebraische Zahlkörper*, J. für Math., v. 169 (1933), pp. 103–7.

HASSE, H. [8]: *Die Normenrestheorie relativ-Abelscher Zahlkörper als Klassenkörpertheorie im Kleinem*, J. für Math., v. 162 (1930), pp. 145–54; [9]: *Die Struktur der R. Brauerschen Algebrenklassengruppe über einem algebraischen Zahlkörper*, Math. Ann., v. 107 (1933), pp. 731–60; [10]: *Klassenkörpertheorie*, mimeographed notes, Marburg, 1932.

HENSEL, K. **[11]**: *Theorie der algebraischen Zahlen*, Leipzig and Berlin, 1908; **[12]**: *Untersuchungen der Zahlen eines algebraischen Körpers für den Bereich eines beliebigen Primteilers*, J. für Math., v. 145 (1915), pp. 92–113; **[13]**: *Die multiplikative Darstellung der algebraischen Zahlen für den Bereich eines beliebigen Primteilers*, J. für Math., v. 146 (1916), pp. 189–215; **[14]**: *Untersuchung der Zahlen eines algebraischen Körpers für eine beliebige Primteilerpotenz als Modul*, ibid., pp. 216–28; **[15]**: *Allgemeine Theorie der Kongruenzklassengruppen und ihrer Invarianten in algebraischen Körpern*, J. für Math., v. 147 (1917), pp. 1–15; **[16]**: *Eine neue Theorie der algebraischen Zahlen*, Math. Zeit., v. 2 (1918), pp. 433–52.

KRASNER, M. **[17]**: *Sur la Primitivité des corps $\mathfrak{P}$-adiques*, Mathematica, v. 13 (1937), pp. 72–191; **[18]**: *Le nombre des surcorps d'un degré*, etc., Comptes Rend., v. 25 (1937), pp. 1026–8; **[19]**: *Le nombre des surcorps primitifs d'un degré*, etc., ibid., v. 206 (1938), pp. 876–8; **[20]**: *Une generalisation de la theorie locale des corps de classes*, loc. cit., pp. 1696–9.

NAKAYAMA, T. **[21]**: *Über die Beziehungen zwischen den Faktorensystemen und der Normklassengruppe eines galoisschen Erweiterungskörpers*, Math. Ann., v. 112 (1936), pp. 85–91.

RICHTER, H. **[22]**: *Über die Lösbarkeit des Einbettungsproblems für Abelsche Zahlkörper*, Math. Ann., v. 112 (1936), pp. 700–26.

SCHILLING, O. F. G. **[23]**: *Arithmetic in fields of formal power series in several variables*, these Annals, vol. 38 (1937), pp. 551–76; **[24]**: *The structure of local class field theory*, Am. J., v. 60 (1938), pp. 75–100; **[25]**: *A generalization of local class field theory*, ibid., pp. 667–704; **[26]**: *A remark on normal extensions*, loc. cit., v. 61 (1939), pp. 388–96; **[27]**: *Regular normal extensions over complete fields*, Trans. Am. Math. Soc., v. 47 (1940), pp. 440–54.

SCHMIDT, F. K. **[28]**: *Zur Klassenkörpertheorie im Kleinen*, J. für Math., v. 162 (1930), pp. 155–68.

WAHLIN, G. E. **[29]**: *The equation $x^l - A = 0(p)$*, J. für Math., v. 145 (1915), pp. 114–38.

ANNALS OF MATHEMATICS
Vol. 41, No. 3, July, 1940

POLYNOMIALS IN SEVERAL VARIABLES[1]

BY RUFUS OLDENBURGER

(Received June 5, 1940)

1. Introduction

Sylvester, Hilbert, Richmond, and Bronowski[2] (I–IV) have concerned themselves with certain very special problems on the expressability of forms of degree p as sums of p^{th} powers of linear forms with coefficients in the complex field of numbers. The class of forms which can be written as such sums for fields other than the complex field is quite limited.

It has been proved (V) that each quadratic form $Q = a_{ij}x_ix_j$ $[a_{ij} = a_{ji}]$ is equivalent under non-singular linear transformations in a given field K to $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2$ where r is the rank of Q. This is equivalent to the statement that Q can be written as a *linear combination* $\lambda_i L_i^2$ *of squares of linearly independent linear forms* $L_1, L_2, \cdots, L_r$ with coefficients in K. Recently the author determined (VI) when a form F of degree p can be written for a field K as a linear combination of p^{th} powers of linearly independent linear forms. The form F is then equivalent to a canonical form

$$\lambda_1 y_1^p + \lambda_2 y_2^p + \cdots + \lambda_r y_r^p \tag{1.1}$$

with diagonal matrix of coefficients. Such forms were termed *non-singular with respect to* K. Not every form F of degree p is non-singular for a field K. We are thus led to consider conditions under which a form can be written for a field K as a linear combination of p^{th} powers of linear forms. Each form which can be so written can be written as

$$a_{ij\cdots k}x_ix_j \cdots x_k \qquad [i, j, \cdots, k = 1, 2, \cdots, n],$$

where the matrix $(a_{ij\cdots k})$ of coefficients is symmetric (that is, the element $a_{ij\cdots k}$ is equal to each element obtained from $a_{ij\cdots k}$ by permutation of the subscripts). It is therefore to be understood throughout this paper that each form has a symmetric *matrix of coefficients*. This restriction is necessary since, for example, the form xy has no symmetric matrix of coefficients for a field with characteristic two.

The basic result of the present paper is that a form F of *degree* p *can be written as a linear combination of* p^{th} *powers of linear forms if the field* K *of coefficients is of order* p *or more.* Call such a linear combination involving a minimum number of terms a *minimal representation* of F for K. Also call the number

[1] Presented to the American Mathematical Society, October 30, 1937, and April 13, 1940.

[2] The reader is referred to papers I–IV listed at the end of this article. Throughout the present paper Roman numerals will be used to designate references in the bibliography.

of terms in a minimal representation the *minimal number* of F for K, and denote it by $m(F)$. Then the number $m(F)$ is *invariant* under non-singular linear transformations on the variables of F, where the coefficients in these transformations are in K. We shall be concerned with the study of the range of the minimal number [in particular for the class of binary forms], the various minimal representations of a form, the minimal numbers of algebraic combinations of forms, and the equivalence of forms.

Evidently the majority of the theorems in the present paper may be stated with slight modification for non-homogeneous polynomials.

2. Definitions

If a form F can be transformed by a non-singular linear transformation with coefficients in a field K to a form with r variables, and F cannot be so transformed into a form with less than r variables, F is said to be a form in r *essential variables,* and we say that "r of the variables of F are essential." We arrange the elements of the symmetric matrix $(a_{ij\cdots k})$ of F in an ordinary 2-way matrix A_2 with i as row index and columns associated with $j, \cdots, k$. A typical column of A_2 is of the form

$$\begin{Vmatrix} a_{1j} & \cdots & k \\ a_{2j} & \cdots & k \\ \vdots & & \\ a_{nj} & \cdots & k \end{Vmatrix}.$$

It can be proved (cf. VII) that *the rank of the matrix A_2 is equal to the number of essential variables in F.* Hence the number of essential variables in a form F can be determined in a finite number of steps. In the case of a quadratic form F the minimal number of F is obviously identical with the rank and number of essential variables in F.

If we can write a form F as

$$\lambda_1 L_1^p + \lambda_2 L_2^p + \cdots + \lambda_\sigma L_\sigma^p, \tag{2.1}$$

with coefficients in a field K, we shall say that (2.1) is a *σ-representation* of F with respect to K.

3. Remarks on forms with symmetric matrices of coefficients

Each form F of degree p in n variables has a symmetric matrix of coefficients if and only if the characteristic of the underlying field K does not divide the coefficients $C(p; r, s, \cdots, t) = p!/r!s! \cdots t!$ in the expansion of $(x_1 + x_2 + \cdots + x_n)^p$. If F and G denote two forms of degree p, and $F = G$ for all values of the variables in a field K, corresponding coefficients in F and G must be equal if and only if K is of order $(p + 1)$ at least. It follows that a form F of degree p in n variables has a *unique symmetric matrix of coefficients if and only if the field K is of order $(p + 1)$ at least, and the characteristic does not divide the*

integers $C(p; r, s, \cdots, t)$. If $n > 2$, this matrix is unique if and only if the characteristic of K is greater than p. If $n = 1$, the matrix of coefficients is a single element and is always unique. For a field for which the symmetric matrices of forms of degree p in n variables are unique, the symmetric matrices of forms of degree p in m, $m < n$, variables are also unique. In the problem of equivalence of forms we shall need these uniqueness notions.

4. Representation of forms

We let N denote the number of distinct terms in the expansion of $(x_1 + x_2 + \cdots + x_n)^p$. The existence of representations is established in the following theorem.

THEOREM 4.1. *For a field of order p or more the minimal number $m(F)$ of a form F of degree p in n essential variables exists, and satisfies the inequalities*

$$n \leqq m(F) \leqq N.$$

That $m(F) \geqq n$ is obvious. We consider the general form $F = a_{ij\cdots m}x_ix_j \cdots x_m$, where $A = (a_{ij\cdots m})$ is symmetric, the p indices $i, j, \cdots, m$ range over $1, 2, \cdots, n$, and the elements of A are indeterminates over K. Now $F \equiv \lambda_1 L_1^p + \lambda_2 L_2^p + \cdots + \lambda_\sigma L_\sigma^p$, if

$$\sum_{\alpha=1}^{\sigma} b_{\alpha i} b_{\alpha j} \cdots b_{\alpha m} \lambda_\alpha = a_{ij\cdots m}, \tag{4.1}$$

where we have written $L_\alpha = b_{\alpha 1}x_1 + b_{\alpha 2}x_2 + \cdots + b_{\alpha n}x_n$. We may write $(x_1 + x_2 + \cdots + x_n)^p$ as a sum

$$\sum_{i=1}^{N} a_i f_i(x),$$

where the a_i are integers, and the f_i are distinct power products of degree p in $x_1, x_2, \cdots, x_n$. We let b_i denote the set of elements $(b_{i1}, b_{i2}, \cdots, b_{in})$ for each value of i in the set $1, 2, \cdots, \sigma$. The system of equations (4.1) with $\sigma = N$ is then equivalent to the set

$$\sum_{\alpha=1}^{N} f_\beta(b_\alpha)\lambda_\alpha = y_\beta \qquad (\beta = 1, 2, \cdots, N), \tag{4.2}$$

where $y_1, y_2, \cdots, y_N$ are equal in some order to the independent elements of A. If we show that the b_α can be chosen in K so that the determinant

$$| D_N | = | f_\beta(b_\alpha) |$$

is not zero, there exist solutions for the λ's in (4.2), and hence an N-representation of each form F of degree p in n essential variables.

We suppose that the f's are labeled so that $f_1(b_1) = b_{11}^p$. Since $f_1(b_1)$ is a non-zero polynomial we can choose b_1 so that $f_1(b_1) \neq 0$. The maximum possible rank of D_N is thus at least 1. We assume that the maximum possible rank of D_N is $M < N$. By rearranging the f_β, if necessary, we may assume the exist-

ence of $b_1, \cdots, b_M$ such that $|D_M| = |f_\beta(b_\alpha)| \neq 0$, $[\alpha, \beta = 1, 2, \cdots, M]$, and that $|D_{M+1}| = |f_\beta(b_\alpha)| = 0$ $[\alpha, \beta = 1, 2, \cdots, M+1]$ for this choice of b_1, $b_2, \cdots, b_M$ and any b_{M+1}. Then the rows of $|D_M|$ are linearly independent, and the row $f_{M+1}(b_1), f_{M+1}(b_2), \cdots, f_{M+1}(b_M)$ is a linear combination of them. Hence there exist λ_i in K such that the function

$$f(x) = f_{M+1}(x) - \sum_{\beta=1}^{M} \lambda_\beta f_\beta(x), \qquad (x) = (x_1, x_2, \cdots, x_n),$$

vanishes for $x = b_1, x = b_2, \cdots, x = b_M$. If K is of order $\geqq (p+1)$ the form $f(x)$ is a non-zero polynomial. Thus there exists a b_{M+1} such that $f(b_{M+1}) \neq 0$ in K. For this choice of b_{M+1} we have

$$|D_{M+1}| = \begin{vmatrix} D_M & U \\ 0 & f(b_{M+1}) \end{vmatrix} = |D_M| \cdot f(b_{M+1}) \neq 0,$$

giving a contradiction. Hence $M = N$, and $m(F) \leqq N$.

If the field K is of order p, then $p = r^i$, where r is a prime. Thus $F = \mu_1 x_1^p + \mu_2 x_2^p + \cdots + \mu_n x_n^p$, whence $m(F) \equiv n$.

If F is non-singular, $m(F)$ attains the lower bound n. It is remarked that the *terms in a minimal representation of F can be grouped to give* $F = F_1 + F_2 + \cdots + F_t$, *where* $F_1, \cdots, F_t$ *are non-singular forms.*

It follows from the proof above that each form F of Theorem 4.1 has an N-representation involving the same linear forms, these representations differing only in the values of the λ's.

COROLLARY 4.1. *For a field K of order p or more, a form F of degree p is a linear combination of p^{th} powers of linear forms.*

The *minimal number depends on the field K* as the following example shows.

We write $Q = 6x^2y^2$. Now $m(Q) = 3$ if K contains an element $\omega \neq 1$ such that $\omega^3 = 1$, and the characteristic of K is not 2 or 3. If K does not contain such an ω and has the same restrictions on the characteristic, Q has the minimal representation

$$\tfrac{1}{2}(x+y)^4 + \tfrac{1}{2}(x-y)^4 - x^4 - y^4,$$

whence $m(Q) = 4$.

5. Remark on existence of representations

The existence of representations of the form F of Theorem 4.1 can be proved more simply than in Section 4 by first showing that $C(p, q)x^{p-q}y^q$ has a $(p+1)$-representation. Assuming that forms of degree p in less than n variables can be written as linear combinations of p^{th} powers of linear forms, one can show that forms of degree p in n variables can be so written. The existence of representations now follows by induction. This type of proof does not yield as much information about the minimal number as the proof given in Section 4, and is therefore omitted.

It can be readily shown that Corollary 4.1 cannot be strengthened to fields of order less than p.

Corollary 4.1 implies that, for each integer r that divides p and field K of order p or more, F can be written for some value of σ as

$$\lambda_1 H_1^r + \lambda_2 H_2^r + \cdots + \lambda_\sigma H_\sigma^r,$$

where the H_i are forms of degree p/r.

6. Uniqueness

The following theorem is valid for each field K for which the symmetric matrix of R is unique.

THEOREM 6.1. *If*

$$R = \lambda_1 L_1^p + \lambda_2 L_2^p + \cdots + \lambda_\sigma L_\sigma^p \tag{6.1}$$

is minimal, the λ*'s are uniquely determined by the* L*'s.*

Let F and $L_1, \cdots, L_\sigma$ be given as in Section 4. The condition that F and R be identical is given by (4.1). If the λ's were not unique, one or more would be arbitrary, whence we could choose some of the λ's equal to zero. In a minimal representation R the λ's are not zero, whence the λ's are unique.

It follows that the matrix M of coefficients of the λ's in (4.1) is of rank σ.

7. Notation

If F has a representation (6.1), *we shall say that* F *has a representation* $[\lambda, B]$, where λ denotes the set $[\lambda_1, \lambda_2, \cdots, \lambda_\sigma]$, and B is the matrix of coefficients of the L's. The form F is non-singular if and only if B is non-singular.

The matrix B in a minimal σ-representation $[\lambda, B]$ of a *binary* form F of degree p, can always be written as

$$B = DC,$$

where

$$D = \left\| \begin{matrix} \rho_1 & & & 0 \\ & \rho_2 & & \\ & & \ddots & \\ 0 & & & \rho_\sigma \end{matrix} \right\|, \quad C = \left\| \begin{matrix} 1 & \beta_1 \\ 1 & \beta_2 \\ \cdot & \cdot \\ 1 & \beta_\sigma \end{matrix} \right\| \quad \text{or} \quad \left\| \begin{matrix} 1 & \beta_1 \\ 1 & \beta_2 \\ \cdot & \cdot \\ 1 & \beta_{\sigma-1} \\ 0 & 1 \end{matrix} \right\|,$$

and the β's in C are distinct. Hence, F has a representation $[\mu, C]$ where $\mu = [\lambda_1\rho_1^p, \lambda_2\rho_2^p, \cdots, \lambda_\sigma\rho_\sigma^p]$. It is no restriction in what follows to treat only minimal representations of binary forms where these representations are of the type $[\mu, C]$.

8. Binary forms

In this section we shall study the range of the minimal number for forms in two essential variables.

We assume throughout this section that the field K is such that each form considered has a unique symmetric matrix of coefficients.

LEMMA 8.1. *The minimal number of xy^q is $q + 1$.*

Equating $\mu_\alpha[x + \beta_\alpha y]^{q+1}$ $[\alpha = 1, 2, \cdots, q + 1]$ to xy^q we obtain a system of $(q + 2)$ non-homogeneous linear equations in the μ's which can be solved for the μ's if and only if there exists a solution for the μ's in the system of homogeneous linear equations

$$(8.1) \qquad \sum_{\alpha=1}^{q+1} \beta^r_\alpha \mu_\alpha = 0 \qquad [r = 0, 1, \cdots, q - 1, q + 1],$$

where the β's are distinct. The determinant of coefficients in (8.1) is a product of the differences between the β's by the sum

$$S_{q+1} = \beta_1 + \beta_2 + \cdots + \beta_{q+1}.$$

It has been proved elsewhere [XII] that there exists a vanishing sum S_{q+1} unless the characteristic of the field K is two, and $(q + 1) = 2$ or $(q + 1) = n - 2$, where n is the order of K. By assumption, if $(q + 1)$ is even, the characteristic of K is not two. If $(q + 1)$ is odd, and the characteristic of K is two, the order of K is even or infinite, whence $(q + 1) \neq n - 2$. Thus in any case there exists a vanishing sum S_{q+1} of distinct β's, and $m(xy^q) \leqq q + 1$.

If we equate xy^q to $\mu_\alpha(x + \beta_\alpha y)^{q+1}$ $[\alpha = 1, 2, \cdots, q]$, or to $\mu_\alpha(x + \beta_\alpha y)^{q+1} + \mu_q y^{q+1}$ $[\alpha = 1, 2, \cdots, q - 1]$, where the β's are distinct, we obtain inconsistent equations in the μ's. By Section 7, $m(xy^q) = q + 1$.

THEOREM 8.1. *For the class C of binary forms F of degree p the minimal number ranges over* 2, 3, $\cdots$, p *or* 2, 3, $\cdots$, $p + 1$.

Since we are dealing with forms in two essential variables $m(F) \geqq 2$.

By Lemma 8.1, we can write

$$(8.2) \qquad xy^{p-1} = \lambda_i L_i^p \qquad [i = 1, 2, \cdots, p].$$

Since the p-representation in (8.2) is minimal any partial sum $\lambda_i L_i^p$ $[i = 1, 2, \cdots, r]$ of (8.2) is minimal. Hence $m(F)$ takes on the values 2, 3, $\cdots$, p for forms F in C. By Theorem 4.1 $m(F) \leqq p + 1$, whence Theorem 8.1 is proved.

9. Overlapping of ranges

The ranges of the minimal number for classes of forms in n essential variables are related.

THEOREM 9.1. *The range of the minimal number for the class C' of forms of degree q in n essential variables overlaps the range of the minimal number for the class C of forms of lower degree p in n essential variables provided the characteristic of the field does not divide $p + 1$, $p + 2$, $\cdots$, q.*

To prove this we shall use the operation of differentiation defined in the usual way for polynomials without reference to limit processes. For a form F in C, we have $m(F) = \sigma$, where $\sigma = n + \rho$. Now $F = \lambda_\alpha L_\alpha^p$ $[\alpha = 1, 2, \cdots, \sigma]$, where the L's are linear forms. Since for each α we have $L_\alpha \neq 0$, we can trans-

form F non-singularly to a form $F' = \lambda_\alpha M_\alpha^p$ in variables $x_1, x_2, \cdots, x_n$, where the coefficient of x_1 in each M_α is not zero. We can write $F' = \lambda'_\alpha Q_\alpha^p$, where the coefficient of x_1 in each linear form Q_α is unity. Then $Q_\alpha = x_1 + N_\alpha$, where N_α is a linear form in $x_2, x_3, \cdots, x_n$. We introduce the form

$$G = \lambda'_\alpha(x_1 + N_\alpha)^{p+1} \qquad [\alpha = 1, 2, \cdots, \sigma]. \tag{9.1}$$

Evidently,

$$(p + 1)F' = \frac{\partial G}{\partial x_1}.$$

By (9.1), we have $m(G) \leqq \sigma$. If $m(G) = \rho < \sigma$,

$$G = \mu_\alpha T_\alpha^{p+1} \qquad [\alpha = 1, 2, \cdots, \sigma - 1],$$

whence $\partial G/\partial x_1$ yields a $(\sigma - 1)$-representation of F'. Thus $m(G) = \sigma$. This proves the theorem.

10. Minimal number of a sum of forms

We shall need the following obvious lemma.

LEMMA 10.1. *Let k be a non-zero element of a given field K. The minimal numbers of forms kF and F, with respect to K, are equal.*

THEOREM 10.1. *The minimal numbers of forms F, G, and $F + G$ for a field K satisfy the inequalities*:

$$m(F) + m(G) \geqq m(F + G) \geqq | m(F) - m(G) |. \tag{10.1}$$

We note that the existence of $m(F)$ and $m(G)$ implies the existence of $m(F + G)$. The left inequality in (10.1) is obvious. Label the forms F and G so that $m(F) \geqq m(G)$. By Lemma 10.1, we have $m(-G) = m(G)$. The left inequality in (10.1) yields

$$m(F + G) + m(-G) \geqq m(F),$$

whence the theorem is proved.

There exist forms F and G for which the equality signs in (10.1) are attained. If F and G involve $m(F)$ and $m(G)$ essential variables respectively, and the variables of F are independent of the variables of G, $m(F + G) = m(F) + m(G)$. If $G = -F$,

$$m(F + G) = | m(F) - m(G) | = 0.$$

COROLLARY 10.1. *Adding a term aL^p, where L is linear, to a form F of degree p at most changes the minimal number by one.*

In the following theorem we give conditions under which a term aL^p can be added to a form F to increase the minimal number of F. The field of coefficients is arbitrary.

THEOREM 10.2. *Let F be a form such that if $\lambda_\alpha L_\alpha^p$ $[\alpha = 1, 2, \cdots, \sigma]$ is a minimal representation of F, the expression $\lambda_\alpha L_\alpha^{p-1}$ $[\alpha = 1, 2, \cdots, \sigma]$ is also minimal.*

If N is a linear form linearly independent of $L_1, L_2, \cdots, L_\sigma$ and $\mu \neq 0$, the minimal number of F is increased by addition of μN^p to F.

A σ-representation of $F + \mu N^p$ can be written as

$$F + \mu N^p = \lambda_\alpha[L_\alpha + \beta_\alpha N]^p, \qquad [\alpha = 1, 2, \cdots, \sigma], \tag{10.2}$$

where $L_\alpha = a_{\alpha j}x_j$ $[j = 1, 2, \cdots, n]$, the variables $x_1, x_2, \cdots, x_n$ being essential variables of F. Then $\lambda_\alpha L_\alpha^p$ $[\alpha = 1, 2, \cdots, \sigma]$ is a minimal representation of F. By assumption $\lambda_\alpha L_\alpha^{p-1}$ is a minimal representation of a form G. We can write G as $a_{ij\cdots m}x_ix_j \cdots x_m$, where the matrix $(a_{ij\cdots m})$ is symmetric. We let M denote the 2-way matrix $(a_{\alpha i}a_{\alpha j} \cdots a_{\alpha m})$ [α not summed] with α as column index, and $(i, j, \cdots, m)$ associated with the rows of M. By Section 6 the matrix M has rank σ. Equating coefficients in (10.2) we obtain

$$\lambda_\alpha \beta_\alpha a_{\alpha i}a_{\alpha j} \cdots a_{\alpha m} = 0 \qquad [\alpha = 1, 2, \cdots, \sigma]. \tag{10.3}$$

Since M has rank σ, the only possible solutions of (10.3) are

$$\lambda_1\beta_1 = \lambda_2\beta_2 = \cdots = \lambda_\sigma\beta_\sigma = 0.$$

Since $\lambda_\alpha L_\alpha^p$ $[\alpha = 1, 2, \cdots, \sigma]$ is minimal, the λ's are not zero, whence $\beta_\alpha = 0$ for each α. Setting the x's equal to zero in (10.2), we obtain $\mu N^p = \lambda_\alpha(\beta_\alpha N)^p$ $[\alpha = 1, 2, \cdots, \sigma]$, whence we have arrived at a contradiction. Thus $m(F + \mu N^p) = m(F) + 1$.

11. Minimal numbers of products of forms

In a polynomial $P_n(\beta) = c_0\beta^n + c_1\beta^{n-1} + \cdots + c_{n-1}\beta + c_n$ we term c_i the i^{th} coefficient. If for a field K there exists a polynomial $P_n(\beta)$ of degree n with distinct zeros and with i^{th} coefficient zero, we say that the field possesses the property $P_{n,i}$. By Theorem 4.1 for any field K $m[C(p + q, p)x^py^q] \leqq p + q + 1$. A stronger result is stated in the following lemma where the field K is understood to be such that the symmetric matrix of x^py^q is unique.

LEMMA 11.1. *The minimal number of $F = x^py^q$, $p \geqq 2$, is less than $(p + q + 1)$ if and only if the field K has the property $P_{p+q,p}$* or *$P_{p+q-1,p-1}$.*

Equating F to

$$R = \mu_\alpha(x + \beta_\alpha y)^{p+q} \qquad [\alpha = 1, 2, \cdots, p + q]$$

we obtain a system of linear equations in the μ's which can be solved for the μ's if and only if there exists a solution for the μ's in the system

$$\sum_{\alpha=1}^{p+q} \beta_\alpha^r \mu_\alpha = 0 \qquad [r = 0, 1, \cdots, q - 1, q + 1, \cdots, p + q], \tag{11.1}$$

where the β's are distinct. The determinant of coefficients of the μ's in (11.1) is a product of the differences between the β's by the coefficient c_p in $P_{p+q}(\beta)$. Hence F has a representation R if and only if K has the property $P_{p+q,p}$. It can be shown that F has a representation $R' + \mu_{p+q}y^{p+q}$, where $R' = \mu_\alpha(x + \beta_\alpha y)^{p+q}$ $[\alpha = 1, 2, \ldots, p + q - 1]$, and the β's are distinct, if and

only if $\partial F/\partial x$ has a representation $\partial R'/\partial x$. In view of Section 7 the lemma is now proved.

LEMMA 11.2. *Let K be a field with characteristic not two, and order at least n. The field K possesses the property $P_{n,i}$ for each odd integer i.*

A pair $(a, -a)$ of elements of K is called a *0-pair*. The pair $(0, 0)$ is termed the *null* 0-pair.

We assume that $n = 2k$, where k is an integer. Since the characteristic of K is odd, the order of K is odd or infinite. There thus exist distinct non-null 0-pairs $(a_1, -a_1)$, $(a_2, -a_2)$, $\cdots$, $(a_k, -a_k)$ in K. In the expansion of

$$(11.2)\quad P(x) = (x + a_1)(x - a_1)(x + a_2)(x - a_2) \cdots (x + a_k)(x - a_k)$$

no terms of odd degree occur.

We let $m = 2k + 1$, where k is an integer. We replace $P(x)$ by $xP(x)$ thus adjoining 0 to the a's in (11.2), whence the lemma is proved.

In the following theorem it is naturally assumed that the field K is such that F, G and FG can be expressed as linear combinations of powers of linear forms.

THEOREM 11.1. *Let K be a field with characteristic not two and let F and G be forms of degrees p and q respectively. The minimal numbers of F, G, and FG with respect to K, satisfy the inequality*:

$$(11.3)\qquad m(FG) \leqq (p + q)m(F)m(G).$$

By Lemma 11.2, the field K possesses at least one of the properties $P_{p+q,p}$ or $P_{p+q-1,p-1}$ when $p \geqq 2$. By Lemma 8.1, and 11.1,

$$m(x^p y^q) \leqq p + q.$$

Since the product of a ρ-representation $\lambda_i L_i^p$ by a σ-representation $\mu_j M_j^q$ is a sum of $\rho\sigma$ terms of the form $\lambda_i \mu_j L_i^p M_j^q$ (i, j not summed), the theorem is proved.

By Lemma 8.1 the equality sign in (11.3) may be attained.

We are now in a position to generalize a well-known theorem on the ranks of reducible quadratic forms.

THEOREM 11.2. *If F is a form of degree p which splits into linear factors with coefficients in K, the minimal number of F with respect to K is not greater than $p!$.*

By Theorem 11.1, if $F = GL$, where L is linear, $m(F) \leqq pm(G)$. Repeated application of this principle yields the proof of the theorem.

12. Equivalence of forms

In what follows "equivalent" will be understood to mean "equivalent under non-singular linear transformations in the given field." It will be assumed throughout this section and sections 13 and 14 that the field is such that the forms have *unique* symmetric matrices.

A minimal representation of a form F in n essential variables is always of the type

$$(12.1)\qquad \sum_{\alpha=1}^{n} \lambda_\alpha L_\alpha^p + \sum_{\alpha=n+1}^{\sigma} \lambda_\alpha \left(\sum_{i=1}^{n} c_{\alpha i} L_i\right)^p,$$

where the L's are linearly independent. The λ's and the c's in (12.1) form a set of elements S. The sets S for a given form F form a class C of sets which we shall term the *minimal class* of F with respect to K.

THEOREM 12.1. *Two forms F and F' are equivalent for a field K, if and only if the minimal classes of F and F' with respect to K are identical.*

A non-singular transformation on the variables of F can be represented by the equalities

$$L_i = R_i \qquad [i = 1, 2, \cdots, n], \tag{12.2}$$

where the R's are linear forms in the new variables. Under a transformation (12.2), the minimal representation (12.1) goes into a minimal representation with the same set S of coefficients.

THEOREM 12.2. *Let $[\lambda, A]$ be a minimal representation of a form F. For a form F' to be equivalent to F it is necessary that F have a minimal representation $[\lambda, A']$ for some matrix A' and the same λ's. If this condition is satisfied, F and F' are equivalent if and only if there exist non-singular matrices U, X such that*

$$A' = UAX, \tag{12.3}$$

where $[\lambda, UA]$ is a representation of F.

If we apply a transformation $x_i = b_{ij}y_j$ to a form F with representation $[\lambda, A]$ we obtain a form F' with representation $[\lambda, AB]$, where B is the matrix (b_{ij}). Since the variables in F are essential, the rank of A is the number of columns of A. Let A' be a matrix of the same dimensions and rank as A. There exists a non-singular matrix U such that $A' = UA$. Theorem 12.2 is now proved.

The problem of equivalence of forms thus divides into the following parts:

1. The determination of all minimal representations $[\lambda, UA]$ of a form with representation $[\lambda, A]$.
2. The solution of (12.3) for a non-singular matrix X.

13. Remark on the equivalence of forms

Let F and F' be forms of degree p in n variables with symmetric matrices $A = (a_{ij\cdots km})$ and $A' = (a'_{qr\cdots tf})$ respectively. We arrange the elements of A in a 2-way matrix A_2 with m as column index so that a typical row of A_2 is

$$\| a_{ij\cdots k1}a_{ij\cdots k2} \cdots a_{ij\cdots kn} \|.$$

In the same way we arrange the elements of A' in a 2-way matrix A'_2 with f as column index so that a typical row of A'_2 is

$$\| a'_{qr\cdots t1}a'_{qr\cdots t2} \cdots a'_{qr\cdots tn} \|.$$

The forms F and F' are equivalent if and only if there exists a non-singular matrix x such that

$$x_D A_2 x = A'_2, \tag{13.1}$$

where x_D is the direct product of $(p - 1)$ matrices identical with x. The matrix x_D is non-singular if x is non-singular, a fact upon which much of the

"rank" theory of the author developed elsewhere (VII) is based. Equation (13.1) is in appearance much like (12.3), but (12.3) is stronger than (13.1) in that the 2-way matrix A of (12.3) has in general less rows than A_2 of (13.1) and the matrix U is often more restricted than x_D.

14. Relations between minimal representations

In the present section we determine the matrix U of (12.3) for non-singular forms. Two sums $\lambda_\alpha L_\alpha^p$ ($\alpha = 1, 2, \cdots, \sigma$) and $\mu_\alpha M_\alpha^p$ ($\alpha = 1, 2, \cdots, \sigma$) are said to be *simply related* if there exist elements $v_1, v_2, \cdots, v_\sigma$ in the given field, and an ordering of the M_α so that

$$M_\alpha = v_\alpha L_\alpha, \qquad \mu_\alpha v_\alpha^p = \lambda_\alpha \qquad [\alpha \text{ not summed}].$$

Evidently, simply related representations are representations of the same form. We shall express this in terms of matrices. We let J denote a permutation matrix (obtained from the identity matrix I by permuting rows of I). For each r there exists a $c(r)$ such that the non-vanishing element in the r^{th} row of J occurs in the $c(r)^{\text{th}}$ column. We choose $\mu_1, \mu_2, \cdots, \mu_\sigma$ and $v_{1,c(1)}, v_{2,c(2)}, \cdots, v_{\sigma,c(\sigma)}$ different from zero in the given field K so that

$$\lambda_{c(r)} = v_{r,c(r)}^p \mu_r. \tag{14.1}$$

We write

$$D = \left\| \begin{matrix} v_{1,c(1)} & 0 & 0 & \cdots & 0 \\ 0 & v_{2,c(2)} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots 0 & v_{\sigma,c(\sigma)} \end{matrix} \right\|.$$

The representation $[\mu, DJA]$ is simply related to $[\lambda, A]$, and each representation simply related to $[\lambda, A]$ is of this form.

THEOREM 14.1. *Let F be a non-singular form of degree at least three. The minimal representations of F are simply related.*

We let $[\lambda, A]$ and $[\mu, B]$ be minimal representations of F. Since F is non-singular, A and B are non-singular. We write $A = (a_{ij})$, $B = (b_{ij})$ and take these to be of order n, whence F involves n essential variables. Equating minimal representations of F we obtain

$$\lambda_\alpha a_{\alpha i} a_{\alpha j} \cdots a_{\alpha m} x_i x_j \cdots x_m = \mu_\alpha b_{\alpha i} b_{\alpha j} \cdots b_{\alpha m} x_i x_j \cdots x_m, \tag{14.2}$$

where the indices range over $1, 2, \cdots, n$. We make the transformation $x_i = c_{is} y_s$, where (c_{is}) is the inverse of A, whence (14.2) yields

$$\lambda_\alpha \delta_{\alpha i} \cdots \delta_{\alpha m} y_i \cdots y_m = \mu_\alpha v_{\alpha i} \cdots v_{\alpha m} y_i \cdots y_m, \tag{14.3}$$

where $(\delta_{\alpha i}) = \cdots = (\delta_{\alpha m}) = I$, I being the Kronecker delta, and $(v_{\alpha i}) = BA^{-1}$. We denote $(v_{\alpha i})$ by Δ. We can write (14.3) as

$$\lambda_\alpha \delta_{\alpha i} \cdots \delta_{\alpha m} y_i \cdots y_m = \mu_\alpha \delta_{\alpha \rho} \cdots \delta_{\alpha \tau} v_{\rho i} \cdots v_{\tau m} y_i \cdots y_m, \tag{14.4}$$

where $(\delta_{\alpha\rho}) = \cdots = (\delta_{\alpha\tau}) = I$. We write

$$z_\rho = v_{\rho i} y_i, \tag{14.5}$$

whence (14.4) is an equality between the forms

$$F = \sum_{\alpha=1}^{n} \lambda_\alpha y_\alpha^p, \qquad G = \sum_{\alpha=1}^{n} \mu_\alpha z_\alpha^p.$$

That is, the form G can be obtained from F by the non-singular linear transformation (14.5). The transformation (14.5) is then such as to bring a form G with diagonal matrix into a form with diagonal matrix. By a theorem of another paper [VI] the only transformations (14.5) with this property are those for which $\Delta = DJ$, where D is diagonal, and J is a permutation matrix. We thus have a $c(r)$ for each r so that $v_{\alpha i} \neq 0$ for $[\alpha, i] = [r, c(r)]$, and zero otherwise. We can now write (14.4) as

$$\sum_{\alpha=1}^{n} \lambda_\alpha y_\alpha^p = \sum_{\alpha=1}^{n} \mu_\alpha v_{\alpha,c(\alpha)}^p y_{c(\alpha)}^p. \tag{14.6}$$

Since (14.6) is an identity in $y_1, y_2, \cdots, y_n$, we may equate coefficients, whence we obtain (14.1). Since $B = \Delta A$, we have $B = DJA$ and $[\mu, B]$ is simply related to $[\lambda, A]$, whence the theorem is proved.

It can be proved that there exist forms with minimal representations not simply related. Thus, for example, not all minimal representations of x^2y are simply related.

15. Multilinear forms

The proofs of the theorems in this section are in general similar to the proofs above for ordinary forms and will be omitted. The theorems of this section are analogues of theorems proved above. The given field of numbers is arbitrary.

THEOREM 15.1. *Let $M = d_{ij\cdots m} x_i y_j \cdots z_m$ be a multilinear form. The form M can be written as*

$$\sum_{\alpha=1}^{\sigma} L_\alpha W_\alpha \cdots N_\alpha, \tag{15.1}$$

where the L's, W's, $\cdots$, N's are respectively linear forms in the x's, y'x, $\cdots$, z's only.

We designate the matrices of the sets of linear forms $[L_\alpha]$, $[W_\alpha]$, $\cdots$, $[N_\alpha]$ by $A, B, \cdots, G$ respectively. If M is written as (15.1), we say that M has a *representation* $[A, B, \cdots, G]$. We write $D = (d_{ij\cdots m})$.

We term the minimum value of σ for which M can be written in the form (15.1) the *factorization rank* of M (and D) with respect to the given field K. That the factorization rank of a form M depends on the field K can be proved by considering binary trilinear forms (VI).

We arrange the elements of the matrix D of M in a 2-way matrix D_2 with i as row index so that a typical column of D_2 has the shape

$$\left\| \begin{matrix} a_{1j\cdots m} \\ a_{2j\cdots m} \\ \vdots \\ a_{nj\cdots m} \end{matrix} \right\| .$$

We term the rank of D_2 the *i-rank* of M. It is equal to the number of linearly independent forms in the set $[L_\alpha]$ of (15.1).

THEOREM 15.2. *Let the maximum of the $i, j, \cdots, m$-ranks of the multilinear forms M of Theorem* 15.1 *be n. Let the degree of M be p. The factorization rank σ of M satisfies the inequalities*

$$n \leqq \sigma \leqq n^{p-1}.$$

If σ in (15.1) is the factorization rank of M the representation (15.1) of M is termed *minimal*.

THEOREM 15.3. *If $[A, B, \cdots, F, G]$ and $[A, B, \cdots, F, G']$ are minimal representations of a multilinear form M, then $G = G'$.*

We term the x-variables in M *essential* if M cannot be brought by a non-singular linear transformation

$$x_i = a_{ij}x'_j$$

into a form with less x'-variables. If the x's in M are essential the i-rank of M equals the number of x's.

THEOREM 15.4. *Let $[A, B, \cdots, G]$ and $[A', B', \cdots, G']$ be minimal representations of multilinear forms M and M' respectively, where all of the variables in M and M' are essential. The forms M and M' are equivalent if and only there exist non-singular matrices $X, \cdots, U, Q, \cdots, S$ such that*

$$XAQ = A', \cdots, UGS = G',$$

where $[XA, \cdots, UG]$ is a representation of M.

We have used "equivalent" in the above theorem in the usual sense that M' can be obtained from M by the independent non-singular linear transformations

$$x_i = a_{ir}x'_r, \quad y_j = b_{js}y'_s, \cdots, z_m = d_{mf}z'_f.$$

Two representations $[A, B, \cdots, F, G]$ and $[A', B', \cdots, F', G']$ of a multilinear form are said to be *simply related* if

$$A' = JUA, \quad B' = JVB, \cdots, F' = JWF, \quad G' = JU^{-1}V^{-1} \cdots W^{-1}G,$$

where J is a permutation matrix, and $U, \cdots, W$ are non-singular diagonal matrices.

THEOREM 15.5. *Let M be a multilinear form of degree p, $p \geqq 4$, given as in Theorem* 15.1, *with all variables essential. Let the $i, j, \cdots, m$-ranks of M be at*

least two. Let M have a representation $[A, B, \cdots, F, G]$, where $B, \cdots, G$ are non-singular. The minimal representations of M are simply related.

A multilinear form M is said to be *non-singular* if M has a representation $[A, B, \cdots, F, G]$, where $A, B, \cdots, F, G$ are ordinary non-singular matrices.

Theorem 15.5 implies the following corollary.

COROLLARY 15.1. *Let M be a non-singular multilinear form of degree at least four. The minimal representations of M are simply related.*

That minimal representations of a multilinear form are not always simply related, and further that Theorem 15.5 cannot be strengthened to include trilinear forms is a consequence of the following example.

Let I denote the identity matrix of order three. We let $B = G = I$, and

$$A = \begin{Vmatrix} 2 & 0 \\ 2 & 0 \\ 0 & 1 \end{Vmatrix}, \quad Q = \begin{Vmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{Vmatrix}, \quad R = T = \begin{Vmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}.$$

For the field K_3, composed of the distinct elements 0, 1, 2, with characteristic 3, $[A, B, G]$, and $[Q, R, T]$ are minimal representations of the same trilinear form. These representations are not simply related.

16. Relations between the factorization rank and minimal number

We let F denote a form with symmetric matrix of coefficients. By applying the well-known polarization process to F we obtain, except for a constant factor, the multilinear form $M_F = a_{ij\cdots m}x_iy_j \cdots z_m$. We shall say that M_F is the *multilinear form associated* with F. The *factorization rank of M_F* with respect to a field K is termed the *factorization rank* of F with respect to K. We denote the factorization rank of F by $f(F)$.

We let $[F]$ denote the class of forms F of degree p in n essential variables. We let K be such that the symmetric matrices of the forms in $[F]$ are unique. The minimum value of the factorization ranks of the forms F in $[F]$ is n. Elsewhere [VI] the author proved that a form F is equivalent to a form $\lambda_1x_1^p + \lambda_2x_2^2 + \cdots + \lambda_nx_n^p$ $[\lambda_1, \lambda_2, \cdots, \lambda_n \neq 0]$ if and only if the associated multilinear form M_F is equivalent to $x_1y_1 \cdots z_1 + \cdots + x_ny_n \cdots z_n$. We thus have the following result.

THEOREM 16.1. *If the minimal number of F in $[F]$ is a minimum for $[F]$, the factorization rank of F is a minimum for $[F]$, and these minimum values are equal; and conversely.*

For the case of quadratic forms the minimal number always takes on its minimum value.

A multilinear form equivalent to $x_1y_1 \cdots z_1 + x_2y_2 \cdots z_2 + \cdots + x_ny_n \cdots z_n$ is *non-singular* in the sense of another paper [X] of the author. Conditions for non-singularity of a multilinear form given in that paper may thus be applied to determine whether or not F is non-singular.

Elsewhere [VI] we obtained directly conditions for the non-singularity of a form F. These conditions are expressed in terms of generalized determinants

involving variable elements, and in terms of the *sub-forms* $F_1, F_2, \cdots, F_n$ of $F = x_i F_i$.

We define the factorization rank $f(A)$ and minimal number $m(A)$ of a symmetric matrix $A = (a_{ij \dots m})$ to be the smallest σ for which

$$A = \left(\sum_{\alpha=1}^{\sigma} b_{\alpha i} c_{\alpha j} \cdots d_{\alpha m}\right) \quad \text{or} \quad \left(\sum_{\alpha=1}^{\sigma} b_{\alpha i} b_{\alpha j} \cdots b_{\alpha m}\right)$$

respectively. That the *factorization rank, and minimal number are not always equal,* at least in the matrix case, follows by considering an example.

We let A be the 6-way matrix $A = (a_{ijklfm})$ of order 2 all of whose elements vanish except those which have exactly one subscript equal to 1. The non-vanishing elements in A are assumed to be equal to 1. If $2, 3 \neq 0$, A is the matrix of $6xy^5$. We let K be the field of order 8 and characteristic 2. For K, the matrix A has a representation

$$\left(\sum_{\alpha=1}^{6} b_{\alpha i} b_{\alpha j} b_{\alpha k} b_{\alpha l} b_{\alpha f} d_{\alpha m}\right),$$

with $b_{11} = b_{21} = \cdots = b_{61} = 1$, and $b_{12}, b_{22}, \cdots, b_{62}$ distinct. Hence $f(A) < 7$. Since there is no vanishing [XII] sum of 6 distinct elements in K, it follows that $m(A) > 6$. Hence $m(A) \neq f(A)$.

17. Note on non-singular forms

In this section the field K is assumed to be such that the forms which occur have unique symmetric matrices. We prove here that if we write a non-singular form F as $x_i F_i$ there exist values ρ_j of x_i so that $\rho_i F_i$ is non-singular. This property simplifies the determination of the non-singularity of a given form. We shall consider first the analogue for multilinear forms, from which the results for ordinary forms are readily derived.

LEMMA 17.1. *Let a multilinear form $G = x_i G_i$ in essential variables x_i, $y_j, \cdots, z_m$ be of degree at least 3, and let G have a representation $[A, B, \cdots, D]$ where $B, \cdots, D$ are non-singular. There exist numbers ρ_i so that $\rho_i G_i$ is non-singular.*

Since $B, \cdots, D$ are non-singular, the indices $j, \cdots, m$ have a common range $1, 2, \cdots, n$, whereas the range $1, 2, \cdots, m$ of i may be different from this. We note that G can be reduced under non-singular linear transformations to a form G' with representation $[A, I, \cdots, I]$, I being the identity matrix of order n. The reduction is performed by operating on the vector spaces $(y_j), \cdots, (z_m)$ with transformations whose matrices are $B^{-1}, \cdots, D^{-1}$ respectively.

We write $G' = x_i G'_i$. Evidently $G' = a_{ti} x_i y'_t \cdots z'_t$ $[i = 1, 2, \cdots, m;$ $t = 1, 2, \cdots, n]$, where $A = (a_{ti})$. Thus $\rho_i G'_i = a_{ti} \rho_i y'_t \cdots z'_t$. If we write $L_t = a_{ti} \rho_i$, the form $\rho_i G'_i$ is simply

$$N = L_1 y'_1 \cdots z'_1 + L_2 y'_2 \cdots z'_2 + \cdots + L_n y'_n \cdots z'_n. \tag{17.1}$$

If $L_1 \equiv 0$, we have $a_{1i} = 0$ for each i, whence G' is free of the variables $y_1', \cdots, z_1'$ contrary to assumption that the variables of G (hence G') are essential. Hence we can choose the ρ's so that (17.1) is non-singular. Applying the transformations $y_t' = b_{tj}y_j, \cdots, z_t' = d_{tm}z_m$ to G' to return to G, the form N transforms covariantly into a non-singular form $\rho_i G_i$.

THEOREM 17.1. *Let F be a non-singular form of degree at least 3. There exists a non-singular linear combination of the sub-forms of F.*

We write F as $a_{ij\ldots m}x_i x_j \cdots x_m$, whence $M_F = a_{ij\ldots m}x_i y_j \cdots z_m$. We introduce the notations $F_i = a_{ij\ldots m}x_j \cdots x_m$, $M_i = a_{ij\ldots m}y_j \cdots z_m$. By Theorem 16.1, the form M_F is non-singular. By Lemma 17.1 there exist numbers ρ_i such that $\rho_i M_i$ is non-singular. Since $\rho_i M_i = M_{\rho_i F_i}$, Theorem 16.1 implies that $\rho_i F_i$ is non-singular when $\rho_i M_i$ is non-singular. The proof of Theorem 17.1 is completed by applying Lemma 17.1 to M_F.

It can be shown that a transformation on $x_1, x_2, \cdots, x_n$ that brings a form

$$\lambda_1 x_1^p + \lambda_2 x_2^p + \cdots + \lambda_n x_n^p \qquad [\lambda_1, \lambda_2, \cdots, \lambda_n \neq 0],\ p \geqq 3,$$

with diagonal matrix into a form with diagonal matrix brings any form in $x_1, x_2, \cdots, x_n$ with diagonal matrix into a like form. From this and Theorem 17.1 a form $F = x_i F_i$ of degree p, $p \geqq 4$, in n variables is *non-singular if and only if*:

a) There exist values ρ_i of x_i so that $\rho_i F_i$ is non-singular.

b) If (a) is satisfied let $\rho_i F_i$ be chosen non-singular and reduce $\rho_i F_i$ by a non-singular linear transformation T to a form with diagonal matrix. Let the range of i be taken so that $\rho_1 \neq 0$. Let the transformation T bring $F_2, \cdots, F_n$ into $F_2', \cdots, F_n'$ respectively. Then $F_2', \cdots, F_n'$ are forms with diagonal matrices.

The present paper suggests various problems whose complete solution is still forthcoming. Among these are the determination of the exact range of the minimal number, the relation between minimal representations, and necessary and sufficient conditions that a representation be minimal. Applications of the theory to classical problems of algebra will appear elsewhere.

BIBLIOGRAPHY

I. J. J. SYLVESTER, Philosophical Magazine, 1851, p. 94.

II. D. HILBERT, Journal de Mathématiques, vol. 4 (1888), pp. 249–286.

III. W. W. RICHMOND, On canonical forms, Quarterly Journal of Mathematics, vol. 33 (1901-02), pp. 331-340.

IV. BRONOWSKI, *The sum of powers of a canonical expression*, Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 69-82.

V. L. E. DICKSON, *On quadratic, Hermitian, and bilinear forms*, Transactions of the American Mathematical Society, vol. 7 (1906), pp. 275-292.

VI. R. OLDENBURGER, *Rational equivalence of a form to a sum of p^{th} powers*, Transactions of the American Mathematical Society, vol. 44 (1938), pp. 219–249.

VII. R. OLDENBURGER, *Composition and rank of n-way matrices and multilinear forms*, Annals of Mathematics, vol. 35 (1934), pp. 622-657.

VIII. CULLIS, *Matrices and determinoids*, vol. 1 (1913), p. 314.

IX. G. W. King, *The indeterminate and composite products of matrices*, Journal of Mathematics and Physics, vol. 13 (1934), p. 445.

X. R. Oldenburger, *Non-singular multilinear forms and certain p-way matrix factorizations*, Transactions of the American Mathematical Society, vol. 39 (1936), pp. 422–455.

XI. Hočevar, *Sur les formes décomposable en facteurs linéaires*, Comptes Rendus, vol. 138 (1904), pp. 745–747.

XII. R. Oldenburger, *Decomposition of elements in abelian groups*, Bulletin of the American Mathematical Society, vol. 45 (1939), pp. 152–158.

Armour Institute of Technology and
the Institute for Advanced Study